Asymmetric transfer of perceptual learning of luminance- and contrast-modulated motion

Alexander A. Petrov   Department of Psychology, Ohio State University, Columbus, OH, USA

Taylor R. Hayes   Department of Psychology, Ohio State University, Columbus, OH, USA

Perceptual learning was used as a tool for studying motion perception. The pattern of transfer of learning of luminance- (LM) and contrast-modulated (CM) motion is diagnostic of how their respective processing pathways are integrated. Twenty observers practiced fine direction discrimination with either additive (LM) or multiplicative (CM) mixtures of a dynamic noise carrier and a radially isotropic texture modulator. The temporal frequency was 10 Hz, speed was 10 deg/s, and duration was 400 ms, with feedback. Group 1 pre-tested CM for 2 blocks, trained LM for 16 blocks, and post-tested CM for 6 blocks during 6 sessions on separate days. In Group 2, the LM and CM roles were reversed. The d' improved almost twofold in both groups. There seemed to be full transfer from CM to LM but no significant transfer from LM to CM. The pattern of post-switch improvement was asymmetric as well—no further learning during the LM post-test versus rapid relearning during the CM post-test. These strong asymmetries suggest a dual-pathway architecture with Fourier channels sensitive only to LM signals and non-Fourier channels sensitive to both LM and CM. We hypothesize that the channels tuned for the same motion direction but different carriers are integrated using a MAX operation.

Introduction

The ability to perceive motion is of crucial importance to ambulatory organisms. A moving object such as a bird gliding across the sky gives rise to a series of luminance changes in the visual image. Such luminance-modulated (LM, first-order, Fourier) motion can be detected by correlation devices such as Reichardt detectors (van Santen & Sperling, 1985) or, equivalently, spatiotemporal energy detectors (Adelson & Bergen, 1985). These models can account for human motion perception of a wide class of stimuli (see Lu & Sperling, 2001b; Papathomas, Rosenthal, & Julesz, 2002, for reviews). However, there are various classes of motion stimuli that are easily perceived by human observers but cannot be detected by first-order mechanisms (e.g., Cavanagh & Mather, 1989; Chubb & Sperling, 1988). These second-order (non-Fourier) stimuli involve isoluminant cues such as contrast, texture, or disparity (see Baker & Mareschal, 2001; Lu & Sperling, 2001b; Wilson, 1999, for reviews). For example, wind blowing through grass gives rise to a series of changes in local contrast across the image, with very little change in luminance.

If spatiotemporal energy detectors cannot detect non-Fourier motion, how is it detected by the brain? The variety of explanations that have been proposed can be organized in two broad classes: single- and dual-pathway theories. The former posit a shared pathway for both first- and second-order motion. The simplest way to make second-order stimuli visible to standard motion analyzers is to add a non-linear pre-processing stage (e.g., Grzywacz, Watamaniuk, & McKee, 1995; Taub, Victor, & Conte, 1997). Other models use a fundamentally different motion detection algorithm based on the ratio of the temporal derivative and the spatial derivative of the image (e.g., Johnston & Clifford, 1995; Johnston, McOwan, & Buxton, 1992; see also Baloch, Grossberg, Mingolla, & Nogueira, 1999). Such gradient-based models can detect various types of second-order motion (e.g., Benton, 2002; Johnston et al., 1992). However, converging evidence from a range of psychophysical (e.g., Edwards & Badcock, 1995; Lu & Sperling, 1995b, 2001b; Nishida, Ledgeway, & Edwards, 1997; Schofield, Ledgeway, & Hutchinson, 2007; Scott-Samuel & Georgeson, 1999; Zanker, 1999), analytic (e.g., Chubb & Sperling, 1988), neurophysiological (e.g., O'Keefe & Movshon, 1998; Zhou & Baker, 1993), neuro-psychological (e.g., Vaina & Soloviev, 2004), and neuro-imaging (e.g., Ashida, Lingnau, Wall, & Smith, 2007) studies indicates that first- and second-order motion CUES are processed by separate pathways, at least initially.

This suggests a dual-pathway architecture. The tuning properties of the first-order (Fourier, "quasi-linear") pathway can be modeled by a bank of linear filters (see Baker

& Mareschal, 2001; Lu & Sperling, 1995b, for reviews). Although it contains non-linearities such as thresholding and contrast gain control (e.g., Heeger, 1992b), they can be accommodated with simple modifications to the linear filter model (see Carandini, Heeger, & Movshon, 1999, for review). Importantly, the Fourier pathway is blind to second-order stimuli, which must be processed separately. The "standard model" for the second pathway is the filter–rectify–filter (FRF) model (Wilson, Ferrera, & Yo, 1992; Zhou & Baker, 1993), also termed "second-order" (Cavanagh & Mather, 1989), "non-Fourier" (Chubb & Sperling, 1988; Wilson, 1999), "back-pocket" (Chubb & Landy, 1991), or "complex channel" model (Graham, Beck, & Sutter, 1992). Although there are differences in the specific details of these various proposals, there is emerging consensus about the filter–rectify–filter core.

Figure 1 is a simplified sketch of the dual-pathway architecture. It is an elaboration of the FRF model of Wilson et al. (1992). It shows three channels tuned for

different directions of motion as indicated by the arrows. The second-order circuits are highlighted in gray. The information flows upward from the input at the bottom. The feedback projections (e.g., Treue & Maunsell, 1996), gain control (e.g., Heeger, 1992b), and internal noise sources (e.g., Lu & Dosher, 2008) are omitted for simplicity. Layer b is a bank of filters tuned for a range of orientations and spatial frequencies (De Valois & De Valois, 1988). The first-order pathway routes the output of these filters directly to the motion extractors in layer e (Adelson & Bergen, 1985; van Santen & Sperling, 1985). The key component of the second-order pathway is the non-linear transformation in layer c. It generates components not present in the Fourier description of the stimulus. The specific form[1] of this non-linearity is not important for our purposes. The rectified signal is then smoothed by a second bank of filters in layer d and routed to the motion extractors in layer e. Within a given direction of motion, there are multiple channels tuned
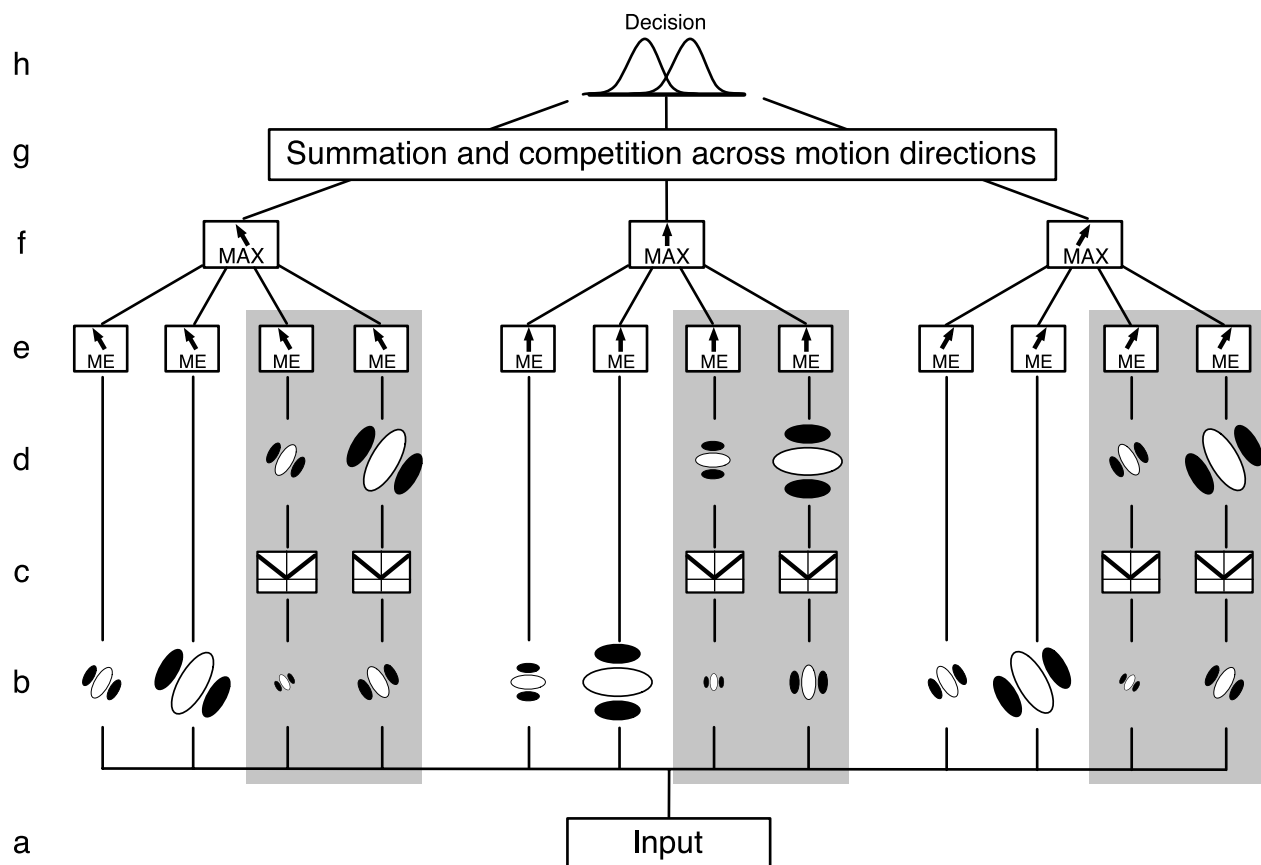

Figure 1. Simplified sketch of a generic dual-pathway architecture for motion processing. It shows three channels tuned for different directions of motion as indicated by the arrows. The second-order circuits are highlighted in gray. Within each motion direction, there are multiple channels tuned for different spatial frequencies. The input in layer a is processed by a bank of early spatial filters (layer b). The first-order pathway routes the filtered signal directly to the motion extractors (ME) in layer e. The second-order pathway includes the filter–rectify–filter cascade in layers b–d. We propose that the two pathways are combined via a MAX operation (layer f) to achieve cue invariance. The information is then integrated across motion directions in layer g, and a discrimination decision is made in layer h. Gain control, internal noise, lateral, and feedback interactions are omitted for simplicity.

for different spatial frequencies (e.g., Cameron, Baker, & Boulton, 1992; Nishida, Ledgeway et al., 1997).

Whereas the Fourier pathway is blind to second-order stimuli, the non-Fourier pathway is not blind to first-order stimuli. The filter–rectify–filter (FRF) cascade is not an impenetrable barrier to many luminance modulations. For example, Edwards and Badcock (1995) performed simulations with a simple FRF model with orthogonal filters (such as those illustrated in the gray boxes in Figure 1). Luminance-defined dots produced a response at the second filter. This is due to the broad spatial frequency content of the dot stimuli and the additional components introduced by the rectification. Luminance-defined plaids are another example (Wilson et al., 1992). The non-linear transformation generates four new grating components: two at the same orientations (but double frequencies) of the plaid gratings and two more at intermediate orientations. The latter two components play a key role in determining the pattern direction of certain ("Type II") plaids in this model (Wilson et al., 1992, p. 81). Thus, not only does the non-Fourier pathway see luminance-modulated stimuli, but it plays an integral role in processing them.

Furthermore, there is psychophysical evidence that the non-Fourier pathway responds to both first- and second-order stimuli. The addition of incoherently moving luminance-defined (LM) dots elevates the coherence threshold of contrast-defined (CM) motion, whereas incoherent CM dots have no effect on the LM threshold (Edwards & Badcock, 1995). Analogous asymmetric influences have been induced with adaptation (Nishida, Ledgeway et al., 1997; Schofield et al., 2007).

The available neurophysiological evidence also suggests the same conclusion. Neuronal responses to various non-Fourier motion stimuli have been recorded in cortical areas MT and MST in monkeys (Albright, 1992; Churan & Ilg, 2001; O'Keefe & Movshon, 1998) and in areas 18 and 19 in cats (e.g., Mareschal & Baker, 1999; Zhou & Baker, 1993; see Baker & Mareschal, 2001, for review). The results showed significant variability along a continuum of directional selectivity to first- and second-order motion. All studies identified a substantial population that is directionally selective to the former but not the latter. Another large population is directionally selective to both. Few neurons were found to respond selectively to second- but not first-order motion. See Appendix A for details.

In sum, in this article we tentatively assume the dual-pathway architecture outlined in Figure 1. The Fourier pathway is sensitive and fast in processing first-order motion but is blind to second-order motion. The non-Fourier pathway can detect both, but its sensitivity and speed are inferior to those of the direct pathway. We eschew the common terminology *second-order pathway* because it invites the misleading notion of a dedicated channel specific to second-order motion.

The outputs of these two pathways are integrated at a later stage into a unified motion percept. It is possible to achieve motion cancellation (rather than transparency) of superimposed oppositely directed luminance and contrast modulations (Lu & Sperling, 1995b, Experiment 4; see also Stoner & Albright, 1992). This suggests that the two pathways ultimately converge. The principles of operation of this integration stage are currently unknown. There is active empirical research on the motion integration in complex stimuli that contain Fourier energy at multiple orientations (e.g., Adelson & Movshon, 1982) and directions (e.g., Qian, Andersen, & Adelson, 1994). Various models have been proposed (see Snowden & Verstraten, 1999, for review). In comparison, the integration of first- and second-order motion has received relatively little attention (Maruya & Nishida, 2010; Stoner & Albright, 1992; Victor & Conte, 1992; Wilson et al., 1992).

The influential model of Wilson et al. accounts for a wealth of behavioral data with plaids (Wilson et al., 1992) and transparent motion stimuli (Kim & Wilson, 1994). The model has a bank of Fourier units with preferred directions spanning 360 deg and a bank of non-Fourier units that also span 360 deg. A third bank of pattern units combine inputs from the component units using a cosine-weighted sum across directions. Translated into the framework of Figure 1, the component and pattern units would occupy layers e and g, respectively. As the weights vary as a function of direction only (Wilson et al., 1992, Equation 8), the Fourier and non-Fourier components *tuned for any particular direction* are integrated via simple (unweighted) summation. In Figure 1, this model would route the signals from layer e directly to layer g, bypassing the MAX operation in layer f.

The present article advances the hypothesis that an additional integration step occurs *within* each motion direction, prior to the global integration *across* directions. We need to introduce some terminology to clarify this distinction. By definition, a set of units (or channels or detectors) is *aligned* when their preferred direction of motion is the same. For example, Figure 1e shows four units that are all tuned for upward motion and are therefore aligned. In addition, a *cue property* is any stimulus property other than the direction of motion. The order of motion (first vs. second) is a cue property and so is spatial frequency. Aligned units can be tuned for different cues as illustrated in Figure 1.

Our working hypothesis is that the units within a given alignment class are integrated to achieve invariance with respect to the cue properties. Following the terminology in the physiological literature (e.g., Albright, 1992; Zhou & Baker, 1993), we will refer to the resulting units as *form-cue invariant*. Our specific hypothesis is that the integration occurs by taking the maximum of the cue-specific signals. This is depicted in Figure 1 by the elements labeled MAX in layer f. The MAX operation is used in

object recognition models (Riesenhuber & Poggio, 1999; Serre, Oliva, & Poggio, 2007) to achieve scale and position invariance. It is superior to averaging because it preserves the sharpness of the tuning curve for the attribute of interest (motion direction in our case) while broadening the tuning for other attributes (Riesenhuber & Poggio, 1999). It can be implemented by neurophysiologically plausible circuits (Yu, Giese, & Poggio, 2002) and accounts well for the response properties of a subclass of complex cells in V1 (Ilan, Ferster, Poggio, & Riesenhuber, 2004) and V4 (Gawne & Martin, 2002). A softened variant of the MAX operation ("softmax") accounts for the integration of end-stopped V1 afferents to MT neurons (Tsui, Hunter, Born, & Pack, 2010).

A related operation is winner-takes-all (WTA, e.g., Yuille & Grzywacz, 1989). MAX collapses the input vector to a single number, whereas WTA transforms it into a sparse vector. The two operations are complementary. MAX preserves the amplitude but not the identity of the "winner neuron," whereas WTA preserves the identity but not the amplitude (Yu et al., 2002). Thus, MAX is useful for achieving form-cue invariance, whereas WTA is useful for selecting a value along some continuum of interest. In motion models in particular, WTA[2] has been proposed for integrating across different directions and velocities (e.g., Nowlan & Sejnowski, 1995; Wilson et al., 1992). In Figure 1, this type of integration is represented in layer g.

Our present proposal is that an additional, MAX integration stage occurs within each set of aligned channels (Figure 1f). The MAX integration promotes form-cue invariance and precedes the WTA competition between the sets. The neural implementation of MAX and WTA (and gain control) is based on the same mechanisms (e.g., feedback, pooled inhibition; Heeger, 1992a, 1992b; Yu et al., 2002; Yuille & Grzywacz, 1989). It seems plausible that the motion system uses both operations in the arrangement in Figure 1. This architecture makes sense from a computational point of view and is also consistent with behavioral data, to which we now turn.

In this article, we use perceptual learning as a tool for studying motion processing. Visual perceptual learning is defined as practice-induced improvement in visual tasks (see Fahle & Poggio, 2002; Fine & Jacobs, 2002; Gilbert, Sigman, & Crist, 2001, for reviews). It has been documented in visual search (e.g., Ahissar & Hochstein, 1997), texture discrimination (e.g., Karni & Sagi, 1991), orientation discrimination (e.g., Dosher & Lu, 1998), Vernier acuity (e.g., Fahle & Edelman, 1993), face identification (e.g., Gold, Bennett, & Sekuler, 1999), and motion detection and discrimination (e.g., Ball & Sekuler, 1987; Huang, Lu, Tjan, Zhou, & Liu, 2008; Law & Gold, 2008; Liu, 1999; Matthews & Welch, 1997; Watanabe, Náñez, & Sasaki, 2001). The learning effects are typically long-lasting and (partially) specific to the particular stimuli used in training (e.g., Ahissar & Hochstein, 1996, 1997,

2004; Ball & Sekuler, 1987; Crist, Kapadia, Westheimer, & Gilbert, 1997; Fahle & Edelman, 1993; Liu, 1999; Matthews & Welch, 1997).

Our present focus is the integration of first- and second-order motion. To our knowledge, there are only two published studies of perceptual learning of non-Fourier motion (Chen, Qiu, Zhang, & Zhou, 2009; Zanker, 1999). Zanker's (1999) pioneering study measured coherence thresholds for coarse direction discrimination (up vs. down) of three types of random dot kinematograms labeled Φ, μ, and θ. All three types were perceived as rectangular objects moving in front of a dynamic noise background. The object consisted of a collection of dots that were in a characteristic relation relative to the observer and to the moving object itself. In Fourier (Φ) motion, the dots moved coherently in the direction of object motion. That is, the dots moved relative to the observer but did not move relative to the object. In drift-balanced motion (μ, Chubb & Sperling, 1988), these relations were reversed: The dots did not move relative to the observer but moved relative to the object. Finally, in motion from motion (θ, Zanker, 1993), the dots moved at a right angle to the direction of object motion. Zanker (1999) referred to Φ as *primary* motion and μ and θ as *secondary* motion. Six experimental groups explored all pairwise combinations of training with one type and testing with another. The main finding was that all three types of training induced perceptual learning but the pattern of transfer was highly asymmetric. Secondary training transferred fully to primary stimuli but primary training did not transfer to secondary stimuli.

This asymmetry adds to the converging evidence against single-pathway theories. The $\theta \rightarrow \Phi$ transfer was interpreted in terms of a hierarchical system in which the output of the motion-from-luminance module is routed into a motion-from-motion module (Zanker, 1993, 1996). It was proposed that Φ training affected only the first module whereas θ training affected both, hence the asymmetrical pattern of transfer.

The hierarchical interpretation does indeed explain why there was *some* transfer from θ to Φ, but it does not explain why there was *full* transfer. The post-switch threshold on the Φ test after θ training (Zanker, 1999, Figure 2, panel b) was statistically indistinguishable from the asymptotic threshold at the end of Φ training (panel a). This is a problem because the Fourier and non-Fourier components move in orthogonal directions in θ stimuli. Thus, practicing up–down θ discrimination exercised the detectors for horizontal Fourier motion. The subsequent Φ test, however, was on vertical Fourier motion. Given that perceptual learning transfers only partially across orthogonal directions (Ball & Sekuler, 1987; Zanker's own Experiment 3), the hierarchical interpretation seems to predict partial $\theta \rightarrow \Phi$ transfer rather than full transfer. There is also a complementary problem. The hierarchical interpretation explains why there was *some* $\Phi \rightarrow \theta$

specificity, but it does not explain why there was *full* specificity. Training with primary stimuli improves the early stage of the motion-from-motion cascade, which should transfer to the subsequent secondary test. However, the post-switch threshold on the θ test after Φ training (Zanker, 1999, Figure 2, panel f) was statistically indistinguishable from the initial θ threshold of untrained observers (panel e). These are between-subject comparisons, and the error bars are inflated by individual differences, which are considerable in perceptual learning. Still, despite its intuitive appeal, the hierarchical interpretation appears less plausible upon closer examination.

There is an alternative interpretation of Zanker's (1999) results in terms of figure–ground segmentation, attention, and/or third-order motion. All three stimulus types involved an object that moved relative to the noise background. Lu and Sperling (1995a, 2001b) proposed a third-order system that computes motion between areas that are marked as figure in successive frames. The so-called inter-attribute motion (Cavanagh, Arguin, & von Grünau, 1989), in which the figure-defining attribute changes from frame to frame, is a striking example of this. Third-order motion is closely tied to attention and it is possible to construct ambiguous displays that change direction depending on which attribute is attended (Cavanagh, 1992; Lu & Sperling, 1995a). It is possible that Zanker's (1999) stimuli engaged the third-order system, even though simulations with a hierarchical model (Zanker, 1996) demonstrated that μ and θ motion can be detected without recourse to attentional or segmentation mechanisms. If μ and θ training did engage these mechanisms, the improvement would transfer fully to Φ motion. On the other hand, the up–down discrimination of a Φ stimulus could also be performed on the basis of local Fourier motion of the dots (which move in the same direction as the object). The observers may have adopted the latter strategy during their initial training with Φ stimuli. However, this strategy is useless for μ and θ stimuli, which explains the complete lack of transfer in this condition. In sum, it is not obvious whether Zanker's (1999) seminal contribution is a study of second- or third-order motion (or a combination thereof).

The experiment of Chen et al. (2009) clearly deals with second-order motion. They measured contrast sensitivity for coarse direction discrimination of drifting luminance- (LM) and contrast-modulated (CM) gratings in the parafovea. The results replicated Zanker's (1999) asymmetric pattern—full[3] transfer of learning from CM to LM and very little transfer from LM to CM. A control group that did the pre- and post-tests but had no training in between provided an estimate of the performance gain attributable to the repeated measurements. That gain on the CM task was similar to the CM gain in the group that trained on LM motion. In other words, there was zero transfer of learning across the LM → CM transition. Another informative feature of the design of Chen et al.

(2009) is that the contrast-sensitivity thresholds were tested at six temporal frequencies ranging from 2 to 30 Hz for LM and from 2 to 21 Hz for CM. The 8-Hz training transferred to all frequencies and led to uniform increments (on a log scale) across the range. This uniformity indicates that the third-order motion system, which is very sensitive to temporal frequency (Lu & Sperling, 1995b), probably plays a negligible role with these stimuli.

Chen et al. (2009) interpreted the asymmetric pattern of transfer in terms of the "different roles played by luminance filters in first- and second-order information processing" (Yifeng Zhou, personal communication, April 8, 2010). The term "luminance filters" refers to the common, early portion of the processing pathways, prior to the rectification in the non-Fourier system (Chen et al., 2009, Figure 6). The corresponding elements in our framework are the early spatial filters in Figure 1b. It is proposed that the LM thresholds are very sensitive to the efficiency of the luminance filters, whereas the CM thresholds are relatively insensitive.

There seem to be two problems with this interpretation. First, the LM and CM stimuli engage early filters of different spatial frequencies (Figure 1). Second, implausible auxiliary assumptions are needed to account for the extreme asymmetry of transfer. Specifically, the complete CM → LM transfer implies that the improvement in LM sensitivity is *entirely* due to improved luminance filters. At the same time, the complete lack of LM → CM transfer implies that improving the same filters has no measurable effect on the CM sensitivity. Thus, one needs to assume that LM and CM training improve the luminance filters equally well, even though the LM threshold depends critically on the efficiency of these filters whereas the CM threshold does not depend at all.

The early-filter interpretation is structurally similar to Zanker's (1999) interpretation and encounters similar problems. Both explain the transfer of learning from non-Fourier to Fourier stimuli by positing plasticity in some shared substrate—the luminance filters (Chen et al., 2009) or the early motion detectors (Zanker, 1999). However, this shared substrate makes it hard to account for the complete lack of transfer in the opposite direction.

The strong asymmetry in the pattern of transfer suggests that a non-linear switch occurs somewhere in the system. Our working hypothesis attributes the switch to the proposed MAX integration of the Fourier and non-Fourier pathways in Figure 1f. This *MAX hypothesis* accounts for the data of Chen et al. (2009) as follows: During LM training, MAX selects the quasi-linear pathway and the plasticity is confined there. LM performance improves, but the improvement cannot transfer to CM because the quasi-linear pathway is blind to contrast modulations. In the other group, CM training strengthens the FRF pathway to the point where it can compete with the quasi-linear pathway for the processing of LM stimuli. During the LM post-test, MAX continues to select the trained FRF

pathway because now it responds more strongly than the untrained quasi-linear pathway. The learning effect thus transfers fully.

The MAX hypothesis makes a critical prediction: There should be no further improvement after the switch to LM following CM training. This is because the same substrate—the FRF pathway—determines the thresholds both before and after the switch. This predicts no drop in performance (that is, full transfer) and no subsequent learning. An extended post-test is necessary to evaluate this prediction. Zanker ([1999]) had such an extended post-test and found no significant improvement in the $\theta \rightarrow \Phi$ group after the switch. This is consistent with the MAX hypothesis, but for his stimuli, it can also be interpreted in terms of third-order motion as discussed above.

The primary goal of the present experiment is to eliminate this ambiguity and test the prediction of the MAX hypothesis with contrast-modulated stimuli and an extended post-test.[4] Arguably, contrast modulation is the paradigmatic second-order signal (e.g., Chubb & Sperling, [1988]; Schofield & Georgeson, [1999]). To minimize the involvement of the third-order motion system, we chose a temporal frequency (10 Hz) high above the cutoff frequency for the third-order system (3–6 Hz). Lu and Sperling ([1995b]) measured the temporal sensitivity functions (TSFs) for various classes of motion stimuli. The sensitivity decreased monotonically with increasing temporal frequency in all cases. The cutoff frequency f_c was defined as the frequency at which the sensitivity drops to one-half of the maximum sensitivity. Luminance- and contrast-modulated stimuli had virtually identical TSFs, with $f_c \approx 12$ Hz (Lu & Sperling, [1995b], Experiment 1). Thus, our 10-Hz stimuli are too fast for the third-order system but within the range of the first- and second-order systems (e.g., Lu & Sperling, [1995b]; Schofield et al., [2007]).

Another goal of the present experiment is to evaluate whether the plasticity site is relatively early or late along the motion-processing stream. As illustrated in Figure 1, the early processing stages are cue specific, whereas the late stages are more integrated and invariant. Consequently, an earlier plasticity site predicts greater specificity of the learning effect, whereas a later site predicts greater transfer (Ahissar & Hochstein, [1997], [2004]). The complete lack of transfer from Fourier to non-Fourier motion in the two published studies suggested an early site (Chen et al., [2009]; Zanker, [1999]). On the other hand, research from our laboratory suggests a late site. A pilot version of the present experiment (Hayes & Petrov, [2009]) found modest transfer of learning from LM to CM motion. In addition, Van Horn and Petrov ([2009]) found complete transfer of learning of LM motion with respect to the spatial frequency (1 vs. 4 cyc/deg) of the carrier. This is consistent with a late plasticity site and in particular with selective reweighting of the connections to the decision stage in Figure 1h. There is growing evidence that such

selective reweighting is a prominent mechanism for perceptual learning (e.g., Dosher & Lu, [1998]; Petrov, Dosher, & Lu, [2005], [2006]). It has been implemented in various models that have accounted successfully for complex behavioral and neurophysiological data (e.g., Law & Gold, [2009]; Lu, Liu, & Dosher, [2010]; Petrov et al., [2005], [2006]; Vaina, Sundareswaran, & Harris, [1995]). In motion direction discrimination in particular, recent neurophysiological studies showed that perceptual learning was accompanied by changes in cortical area LIP but not MT (Law & Gold, [2008]). This strongly suggests plasticity in the connections from MT to LIP.

What is the reason for this discrepancy—apparently early plasticity site in some experiments (Chen et al., [2009]; Zanker, [1999]) and late in others (Petrov et al., [2005]; Van Horn & Petrov, [2009])? We hypothesize that the discrepancy may stem from the difference in their tasks—coarse versus fine discrimination, respectively. In *coarse discrimination* (CD, e.g., Zanker, [1999]), the directional difference between the two stimuli exceeds the full tuning bandwidth of the neuronal populations that represent them. The two populations do not overlap and the most diagnostic neurons are the ones whose tuning curves are centered on the respective stimuli (e.g., Hol & Treue, [2001]). When psychophysical CD sensitivity was compared to information measures from single-cell recordings in monkeys, many individual neurons were as sensitive as the monkey, and the most sensitive neurons were far better than the monkey (Britten, Shadlen, Newsome, & Movshon, [1992]). Readout by simple pooling suffices for CD (Shadlen, Britten, Newsome, & Movshon, [1996]). Coarse discrimination is thus similar to detection. Once the stimulus is detected, its direction is unambiguous and CD thresholds can be statistically indistinguishable from detection thresholds (e.g., Schofield & Georgeson, [1999]). By contrast, in *fine discrimination* (FD, e.g., Van Horn & Petrov, [2009]), the two directions differ only by a few degrees. The two stimulus representations thus overlap significantly and the most diagnostic neurons are the ones whose tuning curve has the steepest slope at the decision boundary (Hol & Treue, [2001]; Regan & Beverley, [1985]). In single-cell recordings in monkeys performing fine discrimination, the individual neurons were far less sensitive than the monkey (Purushothaman & Bradley, [2005]). The optimal FD readout gives highest weight to the flanking neurons and zero weight to the ones centered on the decision boundary (Jazayeri & Movshon, [2006]; Petrov et al., [2005]; Raiguel, Vogels, Mysore, & Orban, [2006]; Seung & Sompolinsky, [1993]).

In sum, the coarse and fine discrimination tasks depend on different neuronal populations and require different readout schemes. Coarse discrimination is similar to detection and can be done through simple pooling, whereas fine discrimination requires sophisticated readout. This invites the hypothesis that CD thresholds may be limited by the sensitivity of the early motion extractors,

whereas FD thresholds may be limited by the precision of the readout. One testable prediction of this hypothesis is that CD training may be more stimulus-specific than FD training. The second goal of the present experiment is to test this prediction. To our knowledge, this is the first study of the perceptual learning of fine discrimination of non-Fourier motion.

Experiment

Ten observers were pre-tested on a fine motion direction discrimination task with luminance-modulated (LM, first-order) stimuli. Then, they practiced the same task with contrast-modulated (CM, second-order) stimuli for 4 sessions. Finally, they were post-tested with LM stimuli for another 1.5 sessions. Ten more observers followed the complementary protocol: CM pre-test, LM training, extended CM post-test.

The experiment is designed to test the two hypotheses introduced above. The extended post-test allows us to test whether additional learning occurs after the switch. The MAX hypothesis predicts that no such learning should occur in the CM → LM group. The pre-test provides a within-subject baseline for measuring the transfer of learning. The late plasticity hypothesis predicts non-zero transfer in both groups. The high-frequency (10 Hz) stimuli are chosen to minimize the involvement of the third-order motion system and the fine discrimination task promotes late plasticity.

Many studies of motion perceptual learning involve fine same–different discriminations (e.g., Ball & Sekuler, 1987; Liu, 1999). The same–different procedure is eschewed here because the multiplicity of possible decision strategies complicates the interpretation of the data (Petrov, 2009). To discourage sequential comparisons across trials, the experimental blocks involve four directions instead of just two. This also allows us to track the performance at two separation levels ($\Delta = 11°$ and $\Delta = 8°$).

Methods

Stimuli

Each stimulus was a mixture of a moving modulator $M(x, y, t)$ and a dynamic noise carrier $N(x, y, t)$ (Figure 2). To generate a luminance-modulated (LM, first-order) stimulus, these ingredients were combined additively according to Equation 1, where $\alpha = 0.13$ is a signal-strength parameter and $C(x, y, t)$ is the relative luminance of pixel (x, y) in frame t. To generate a contrast-modulated (CM, second-order) stimulus, the same two ingredients were combined multiplicatively according to Equation 2, with signal strength $\beta = 0.90$. The strength parameters were estimated in a pilot study[5] so that the initial difficulty
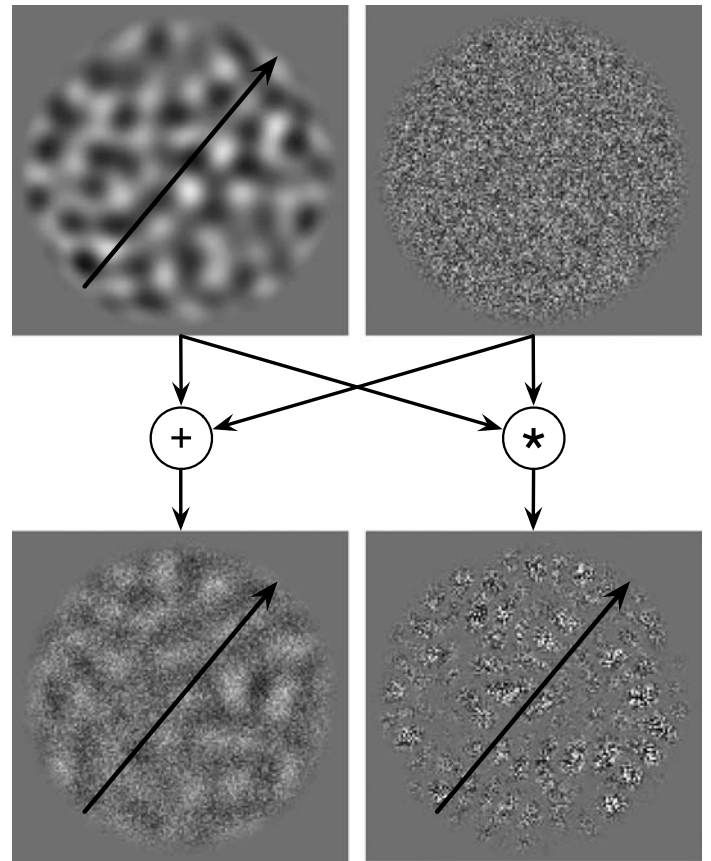

Figure 2. The contrast-modulated stimulus (second-order, bottom right) is a multiplicative mixture of a band-pass texture (top left) and dynamic noise (top right). The luminance-modulated stimulus (bottom left) is an additive mixture of the same texture and noise. The additive signal-to-noise ratio is exaggerated for illustrative purposes. The modulating texture moves from frame to frame in the direction of the arrow.

of the two stimulus types was approximately equal (Ahissar & Hochstein, 1997):

$$C(x, y, t) = [\alpha M(x, y, t) + (1 - \alpha)N(x, y, t)], \quad (1)$$

$$C(x, y, t) = [\beta M(x, y, t) + (1 - \beta)]N(x, y, t), \quad (2)$$

$$L(x, y, t) = L_0 + C(x, y, t)(L_{max} - L_0). \quad (3)$$

All stimuli were generated in real time using Matlab (The MathWorks, 1999) and the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997). The presentation duration was 400 ms (30 frames at 75 Hz). Each noise frame was a matrix of square tiles with size 2 pixels $\approx$ 2.78 arcmin.

The intensity of each tile in each frame was independently drawn from a Gaussian distribution with mean 0 and standard deviation 0.30, clipped between -1.0 and 1.0.

The commonly used (e.g., Chubb & Sperling, 1988) drifting grating modulator is not appropriate for our fine discrimination task because the motion direction is confounded with the static orientation of the grating. A radially isotropic modulator is necessary. The modulator also needs to have a well-defined temporal and spatial frequency. We used filtered-noise textures whose power spectrum peaked at 1 cycle per degree at all orientations (Figure 2, top left). The filter had a Gaussian cross-section with 1 octave full-width at half-height in the Fourier domain. On each trial, the filter was applied to a fresh 512×512 sample of iid Gaussian noise. The mean of the filtered patch was exactly 0 and its expected standard deviation was 0.30. Thirty consecutive frames $M(x, y, t)$ were cut out of this large patch by sliding (and interpolating linearly) a square window in the desired direction. At a speed of 10 degrees per second, the dominant temporal frequency was 10 Hz—well above the cutoff frequency of the third-order system (Lu & Sperling, 1995b).

The mixture of signal and noise was presented on the monitor so that the neutral point $C = 0$ mapped onto the gray background luminance $L_0 = 60$ cd/m^2 (Equation 3). Finally, a circular mask was applied. The mask had a semitransparent linear ramp with an inner rim of 5.5 and an outer rim of 6.5 degrees in diameter.

Apparatus

The movies were presented on a 21″ NEC Accusync 120 color CRT driven by the ATI Radeon HD2600 Pro video card of a 2.66-GHz Intel iMac computer. The monitor was the only light source in the room and was viewed binocularly with the natural pupil from a chin rest located $\approx$92 cm away. At that distance, 1 degree of visual angle spanned $\approx$43 pixels (1024 $\times$ 768 resolution).

The equipment was calibrated to minimize first-order artifacts in the second-order stimuli (Lu & Sperling, 2001a; Smith & Ledgeway, 1997). Three known sources of display non-linearities were addressed: monitor gamma non-linearity, adjacent pixel non-linearity, and noise clumping. The monitor gamma function was estimated via a psychophysical matching procedure (cf., Colombo & Derrington, 2001) and was verified with a Minolta 1° luminance meter. A software lookup table defined 255 evenly spaced luminance levels between $L_{min} = 2$ cd/m^2 and $L_{max} = 118$ cd/m^2. The adjacent pixel non-linearity (APNL) is a hardware problem that reduces the mean luminance of high-contrast transitions along the same video scan line (Klein, Hu, & Carney, 1996). APNL is most severe for static vertical square-wave carriers at high frequencies. It is reduced to negligible levels by our dynamic Gaussian carrier (Smith & Ledgeway, 1997). This was verified by direct measurements with the

photometer. At maximum contrast, the average luminance of the carrier differed from the uniform background luminance by less than 0.5%. Most of this discrepancy is predicted by the natural variability of the noise. Noise clumping can introduce first-order artifacts when local patches of noise have non-zero means (Smith & Ledgeway, 1997). Each blob-like region of our texture covers $\approx$50 noise tiles. At that grain size, the local anisotropies are negligible (Schofield & Georgeson, 1999, Appendix A). Further reduction of the grain size would be counter-productive because of increased APNL. Concerns about non-linearities in the early visual system are addressed in Appendix B.

Observers

Twenty students with normal or corrected-to-normal vision participated in the study. They were paid $6 per hour plus a bonus contingent upon their accuracy.

Task

The fine direction discrimination task was defined with respect to a reference direction θ that was either $-55°$ or $35°$ from vertical. The reference direction was fixed for each participant (except for a short demo block on Day 1) and was counterbalanced between participants. Each block began with two demo trials that explicitly indicated the reference by drawing a line on the screen. The actual direction of motion could take four possible values: $(\theta - 5.5)$, $(\theta - 4)$, $(\theta + 4)$, and $(\theta + 5.5)$. The instructions designated the first two as "counterclockwise" and the last two as "clockwise." The observers indicated their binary response by pressing F or J on the keyboard. Each block presented all four directions with equal frequencies and in random order. This design encouraged absolute discrimination and discouraged comparison with the previous stimulus.

Procedure

The timeline of each trial is illustrated in Figure 3. The trial began with a brief alert sound followed by a 500-ms delay. The motion stimulus was presented for 400 ms, after which the fixation dot reappeared and the behavioral response was recorded. Auditory feedback (an unpleasant beep) was given after incorrect responses. The observers scored one bonus point for each correct response and lost one point for each incorrect response. The cumulative bonus was visible at all times and provided visual feedback. Response times were measured from the stimulus onset. The stimulus was not terminated if the observer responded during the motion presentation. To discourage blind guessing, responses faster than 250 ms triggered a "slow down" message that stayed on the screen
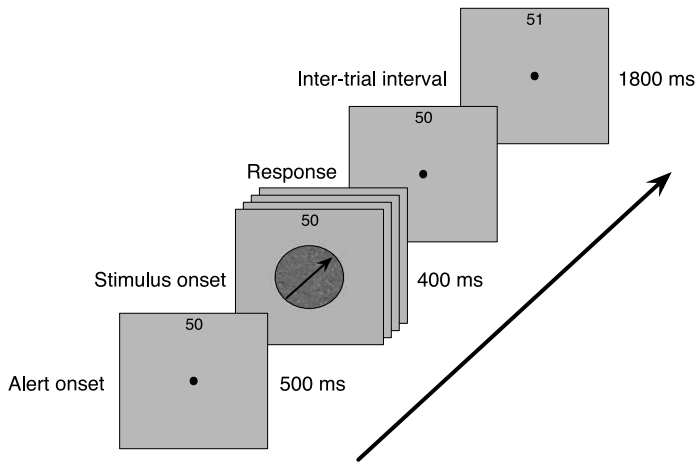
Figure 3. Display layout and trial sequence. The motion stimulus was presented in the middle of the screen for 400 ms. The bonus points (50 in this example) were displayed above the fixation dot and were updated depending on the response.

for 1500 ms. Such fast trials were repeated at the end of the block, as were trials in which an invalid key was pressed.[6] The next stimulus was generated during the inter-trial interval, which took approximately 2000 ms.

The first session began with a verbal instruction and a physical demonstration using a printed texture that moved behind a circular aperture in a piece of cardboard. This was followed by a set of demo trials that were repeated until the experimenter was confident that the participant understood the task. Then, the reference direction was switched and the pre-test began.

Experimental design

The participants were randomly assigned to two experimental groups. Group 1 trained on luminance modulation and transferred to contrast modulation. In Group 2, the roles were reversed. Each observer completed 6 sessions on separate days. Each session consisted of 4 blocks of 208 trials each, for a total of 24 blocks and 4992 trials. Blocks 1 and 2 pre-tested CM in Group 1 and LM in Group 2. The main training—LM in Group 1 and CM in Group 2—began on block 3 and extended through block 18. Finally, blocks 19 to 24 post-tested transfer to CM in Group 1 and LM in Group 2.

Data analysis

Two discriminability values (d') were calculated in each block: for the easy ($\Delta = 11°$) and difficult ($\Delta = 8°$) pair of stimuli. The data from each participant were thus reduced to two parallel d' profiles across the 24 blocks. To smooth out the noise and further reduce the data, a non-linear regression (Equations 4 and 5) was performed on the group-averaged d' profiles.

We assume an exponential law of practice (Heathcote, Brown, & Mewhort, 2000). The standard parameterization in terms of an infinite asymptote is mathematically elegant but leads to numerically unstable parameter estimates. It is better to use the following functional form, which is algebraically equivalent to the standard exponential equation but avoids extrapolations to infinity:

$$f_{m,n}(t; d_m, d_n, \tau) = d_m \frac{e^{-t/\tau} - e^{-n/\tau}}{e^{-m/\tau} - e^{-n/\tau}} + d_n \frac{e^{-m/\tau} - e^{-t/\tau}}{e^{-m/\tau} - e^{-n/\tau}}. \tag{4}$$

Despite its formidable appearance, Equation 4 is very easy to interpret. The parameters d_m and d_n fix the performance levels at two reference times m and n, as illustrated in Figure 4. The reference times are chosen *a priori* on the basis of the experimental design. The performance $f_{m,n}(t)$ at time t is interpolated between these two fixed points according to an exponential law with time constant τ. Equation 4 is translation invariant—that is, $f_{m,n}(t) = f_{m+c,n+c}(t + c)$ for any constant c:

$$\hat{d}(b) = \begin{cases} d_1 & \text{if } b = 1, \\[1em] d_2 & \text{if } b = 2, \\[1em] f_{3,18}(b; d_3, d_{18}, \tau_1/208) & \text{if } 3 \leq b \leq 18, \\[1em] f_{19,24}(b; d_{19}, d_{24}, \tau_2/208) & \text{if } 19 \leq b \leq 24. \end{cases} \tag{5}$$

The piecemeal Equation 5 describes a given d' profile in an economical and theoretically neutral way. It posits separate exponential curves for the pre-test, training, and post-test segments. The d' within each segment is assumed
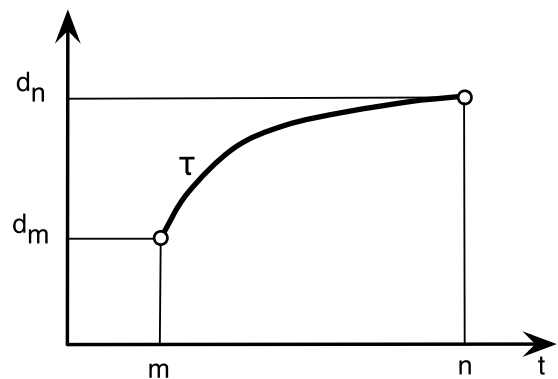

Figure 4. Illustration of the exponential Equation 4. The parameters d_m and d_n fix the performance levels at reference times m and n. The time-constant parameter τ controls the degree of non-linearity.

to improve according to an exponential law, but no assumptions are made about how d' transfers between segments. Equation 5 has 8 free parameters. Six of them are subscripted by block number—d_1, d_2, d_3, d_{18}, d_{19}, and d_{24}—and describe the d' at the beginning and end of their respective segments. The remaining two parameters estimate the time constants during training (τ_1) and post-test (τ_2). The division by 208 converts the time units from blocks to trials to facilitate comparison with published data sets.

It will be important to quantify the transfer of learning from training to post-test. In our experimental protocol, the critical switch occurs after block 18 (trial 3744). This leads to the following *transfer index* (TI):

$$\text{TI} = \frac{d_{19} - d_1}{d_{18} - d_1}. \tag{6}$$

The transfer index is based on the regression parameters of Equation 5: d_1 is the initial d' at pre-test, d_{18} is the d' at the end of training, and d_{19} is the post-test d' immediately after the switch. In words, the transfer index measures the improvement on the test stimuli relative to the total improvement. It complements the *specificity index* of Ahissar and Hochstein (1997). The comparison is done within group, which increases statistical power but requires that the two stimulus classes yield comparable d's. This requirement is met in our data, as we shall see.[7]

Results and discussion

Descriptive statistics

Figure 5 plots the group-averaged d' profiles for the two difficulty levels ($\Delta = 11°$ and $8°$) in the two groups. When fitted to these 4 profiles independently, Equation 5 accounts for over 92% of the variance with 32 free parameters ($R^2 = 0.922$, RMSE = 0.089). Hierarchical regression techniques identify a much more parsimonious model with 11 parameters ($R^2 = 0.913$, RMSE = 0.093). The estimates of these 11 parameters are listed in Table 1. They provide a succinct, high-level description of the data.

The hierarchical regression is described in detail in Appendix C. Briefly, the time constants are poorly constrained by the data and hence common values can be used for all 4 profiles. In addition, the d' on the easy stimulus pair is proportional to that on the difficult pair. Therefore, Table 1 lists the d_i parameters for the difficult ($\Delta = 8°$) profiles only. The corresponding "easy" values can be obtained by multiplying by $k = 1.35$. These regularities eliminate 17 parameters from the saturated model with negligible reduction of the goodness of fit ($F(17,63) < 1$, n.s.).

Recall that noise was added to the luminance-modulated stimuli to equilibrate the initial performance in the two groups. The data indicate that this manipulation is successful. The pre-test d' does not differ significantly between the two types of modulation and the parameters d_1, d_2, and d_3 in Table 1 are common for both groups ($F(3,56) < 1$, n.s.). Thus, the differences at post-test cannot be attributed to differences in task difficulty (Ahissar & Hochstein, 1997).

Learning effects

Figure 5 shows a clear learning trend during the training period (blocks 3–18) in all conditions. The difference $d_{18} - d_3$ estimates the learning effect. It is 0.57 ± 0.23 for Group 1 and 0.44 ± 0.23 for Group 2 ($\pm 90\%$ bootstrap confidence intervals, see Appendix C for details). The small advantage for Group 1 is not statistically significant given the inter-group variability ($z = 0.59$, n.s.). The approximate equality of the total learning effects in the two groups allows us to calculate transfer indices according to Equation 6.
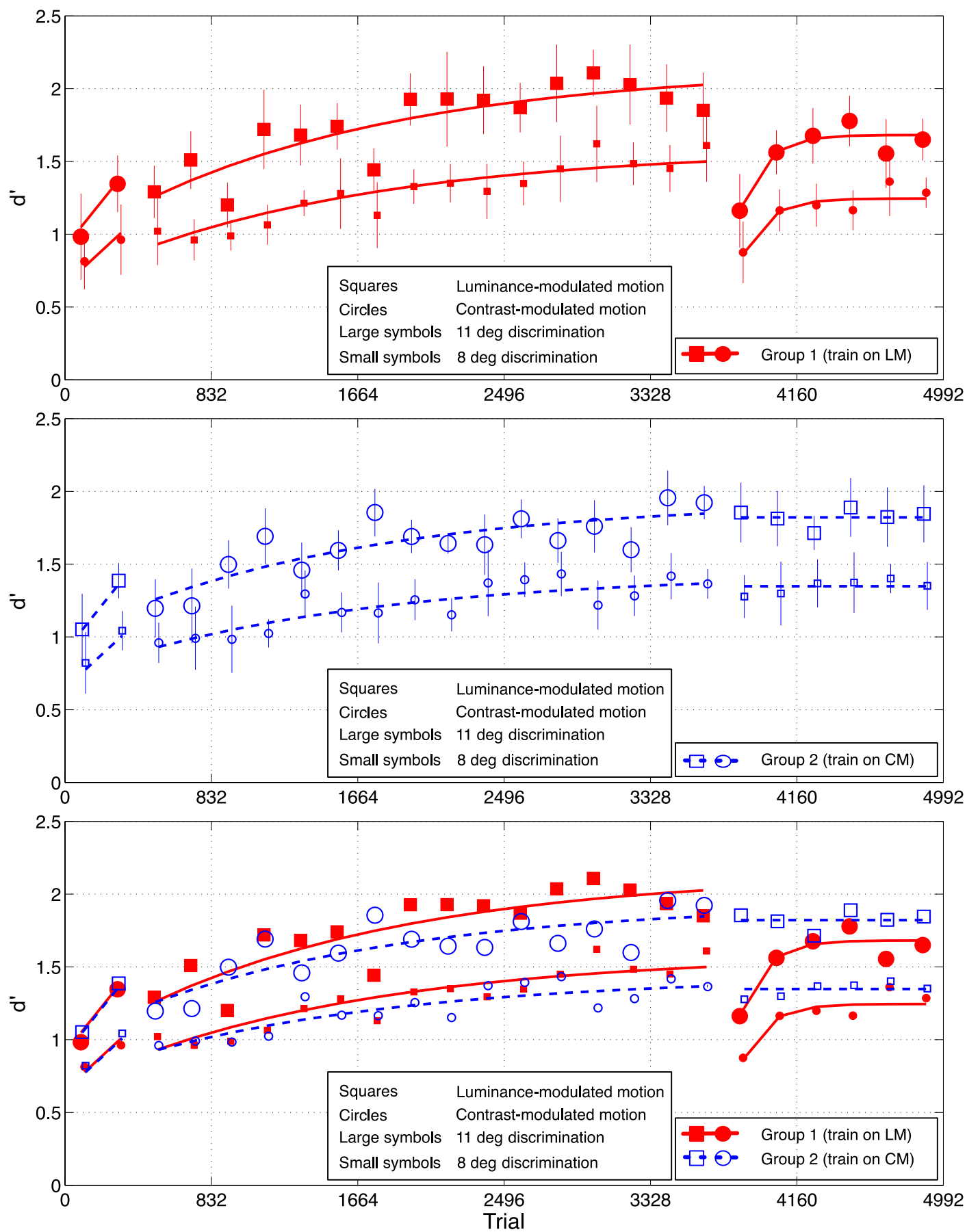
Asymmetric pattern of transfer

Although the two groups improved along very similar trajectories, their post-switch performance was strikingly different. LM training did not transfer to CM (Group 1, Figure 5, top), whereas CM training transferred fully to LM (Group 2, middle). The transfer indices in Table 2 quantify this asymmetry. There is hardly any transfer (12%) in Group 1 and full transfer (97%) in Group 2. The transfer indices are hard to estimate reliably because individual observers improve by different amounts, which destabilize the denominator in Equation 6. This numerical instability is evident in the relatively wide bootstrap confidence intervals in Table 2 (see Appendix C for details). Still, the evidence for a strongly asymmetric pattern of transfer is overwhelming. The regression coefficients in Table 1, the group-averaged d' profiles in Figure 5, and the individual learning curves for most observers in their respective groups all point to the same conclusion.

In sum, the pattern of transfer seems as asymmetric in our fine discrimination task as it is in previous studies with coarse discrimination (Chen et al., 2009; Zanker, 1999). The strong asymmetry suggests that learning takes place at two (or more) separate plasticity sites regardless of the task. This in turn suggests that the brain circuits for first- and second-order motion cannot overlap completely.

Asymmetric pattern of post-switch improvement

This brings us to the extended post-test and the critical prediction of the MAX hypothesis. Recall that it predicts complete lack of learning after the switch to LM stimuli in Group 2. The data in Figure 5 (middle panel) confirm this prediction. There is no significant improvement across

Parameter		Estimate
Beginning pre-test level, both groups	d_1	0.78 ± 0.10
Final pre-test level, both groups	d_2	1.01 ± 0.08
Beginning training level, both groups	d_3	0.93 ± 0.07
Final training level, Group 1	$d_{18,1}$	1.50 ± 0.19
Final training level, Group 2	$d_{18,2}$	1.37 ± 0.12
Beginning post-test level, Group 1	$d_{19,1}$	0.86 ± 0.10
Final post-test level, Group 1	$d_{24,1}$	1.24 ± 0.10
Post-test level throughout, Group 2	$d_{19,2}$	1.35 ± 0.13
Easy/hard proportionality, both groups	k	1.35 ± 0.02
Training time constant, both groups	τ_1	1562 ± 608
Post-test time constant, Group 1	τ_2	138 ± 65

Table 1. Regression model parameters (± 1 bootstrap standard deviation) for the data in Figure 5. The second column gives the corresponding symbol in Equation 5. The d parameters are in d' units and their subscripts correspond to block numbers.

blocks 19–24 (trials 2745–4992) in this group. The hierarchical non-linear regression in Appendix C verifies this. Constraining the parameters of Equation 5 so that $d_{19,2} = d_{24,2}$ does not reduce the fit significantly ($F(1,83) < 1$, n.s.; the unconstrained estimates are $d_{19,2} = 1.33$ and $d_{24,2} = 1.35$). Table 1 therefore lists only one post-test parameter for Group 2.

The seamless transition from CM training to LM post-test suggests that the same pathway processes both types of stimuli. This is consistent with the MAX hypothesis because the CM training may have strengthened the filter–rectify–filter pathway to the point that it outperforms the quasi-linear pathway on LM stimuli. Apparently, the MAX integration stage in Figure 1f continues to select the FRF pathway after the switch and the d' remains at asymptote. No further improvement occurs because 4 sessions of CM training have saturated nearly all plasticity possible for the neural substrate in the FRF pathway.

The post-switch dynamics in the other experimental group is very different, but it too is consistent with the MAX hypothesis. Consider the beginning of LM training (block 3, trial 417 in Figure 5, top). Although the LM stimuli activate both pathways, the FRF pathway apparently is not strong enough to compete with the quasi-linear pathway during the early training. Therefore, it is the quasi-linear pathway that controls behavior and reaps

Figure 5. Group-averaged d' learning curves (error bars are 90% confidence intervals within subjects). Each block consists of 104 easy (large symbols) and 104 difficult (small symbols) trials. The discontinuities in the lines mark switches from luminance- (LM) to contrast-modulated (CM) motion and vice versa. Pre-test on blocks 1–2 (trials 1–416), training on blocks 3–18 (trials 417–3744), extended post-test on blocks 19–24 (trials 3745–4992). (Top) LM training does not transfer to CM (Group 1, 10 observers). (Middle) CM training transfers fully to LM (Group 2, 10 more observers). (Bottom) The same data are superimposed for comparison.

	Group 1 LM → CM	Group 2 CM → LM
Regression model	0.12 (−0.11, 0.31)	0.97 (0.79, 1.24)
Raw data ($\Delta = 11°$)	0.19 (−0.05, 0.39)	1.04 (0.80, 1.40)
Raw data ($\Delta = 8°$)	0.09 (−0.18, 0.32)	0.85 (0.53, 1.14)

Table 2. Transfer indices for the two experimental groups (80% bootstrap confidence intervals in parentheses). Calculated from the regression coefficients in Table 1 and the group-averaged d' for luminance-modulated (LM) and contrast-modulated (CM) motion direction discriminations at two difficulty levels.

the benefit of practice. The LM d' improves, but the learning effect cannot transfer to CM because the MAX operator switches to the FRF pathway at post-test. It is this abrupt non-linear switch that accounts for the abrupt drop in the overt performance.

The CM performance immediately after the switch ($d_{19,1} = 0.86 \pm 0.10$, Table 1) is not significantly different from its pre-test level ($d_1 = 0.78 \pm 0.10$). The statistical power of this comparison is limited by the variability among the individual observers. The confidence intervals in Table 2 suggest that the LM → CM transfer index can be as low as zero (or even slightly negative) or as high as 39% in the group average. A larger sample size is needed to reduce the uncertainty. We have data from 12 additional observers[8] in a related experiment with the same stimuli (Hayes & Petrov, 2009). The confidence interval for the transfer index in this sample was from 30% to 72%. Combining the two data sets, it seems that the population mean is probably between 15% and 45%. In conclusion, the amount of transfer of LM training to CM test is highly variable among individuals and seems quite low on average.

General discussion

Our goal in this article was to use perceptual learning as a tool for studying motion perception. We obtained significant learning effects (almost twofold improvement in d', Table 1) for both luminance- and contrast-modulated motion in a fine direction discrimination task. There was a striking asymmetry in the pattern of transfer of learning across the two stimulus types—full transfer from CM to LM (transfer index TI = 0.97, Table 2) but no significant transfer from LM to CM (TI = 0.12). The pattern of post-switch improvement was asymmetric as well. There was rapid and pronounced improvement during the CM post-test in Group 1 (Figure 5, top) but no improvement during the LM post-test in Group 2 (middle).

These marked asymmetries seem incompatible with the single-pathway theory of visual motion processing (e.g., Grzywacz et al., 1995; Johnston & Clifford, 1995; Taub

et al., 1997). This theory predicts high and approximately symmetric transfer. It may be possible to accommodate some modest degree of asymmetry in a single-pathway model (e.g., Johnston et al., 1992), although we are not aware of any research along those lines. It seems impossible, however, to accommodate the massive asymmetry in our data set. The lack of transfer from LM to CM stimuli establishes the existence of mechanisms that are critical for CM processing but are not improved during LM training. Our result thus adds to the growing evidence against the single-pathway theory (e.g., Ashida et al., 2007; Edwards & Badcock, 1995; Lu & Sperling, 2001b; Nishida, Ledgeway et al., 1997; Schofield et al., 2007; Vaina & Soloviev, 2004; Zhou & Baker, 1993).

The strongly asymmetric pattern of transfer of perceptual learning of Fourier and non-Fourier motion has now been replicated in three independent laboratories with different tasks and stimuli: (a) coherence thresholds for coarse direction discrimination of motion-from-motion kinematograms in the fovea (Zanker, 1999), (b) contrast thresholds for coarse direction discrimination of LM and CM gratings in the parafovea (Chen et al., 2009), and (c) d' for fine direction discrimination of LM and CM textures in the fovea (Hayes & Petrov, 2009, and the present study).

What does this robust result teach us about the organization of the motion system? The overall empirical pattern contains four interlocking components: (a) *some* specificity of learning in the LM → CM direction, (b) *some* CM → LM transfer, (c) apparently *full* CM → LM transfer, and (d) apparently *full* LM → CM specificity. Each of these provides useful constraints on the theory. Let us consider them in turn.

The specificity of learning in the LM→CM direction is the easiest to interpret. We agree with Chen et al. (2009) and Zanker (1999) that this specificity suggests a dual-pathway architecture in which second-order processing depends on some mechanisms that are not improved during first-order training. In the tentative sketch in Figure 1, these are the filter–rectify–filter (FRF) channels.

On the other hand, the transfer of learning in the CM → LM direction suggests an overlap in the mechanisms engaged in the two conditions. There are three different proposals about the specific nature of this overlap. Zanker (1999) proposed a hierarchical system in which a motion-from-luminance module feeds into a motion-from-motion module. The transfer of learning was attributed to plasticity in the motion-from-luminance module, which was engaged by all stimulus types in Zanker's experiment. For contrast-modulated motion, Chen et al. (2009) proposed an analogous scheme in which the early luminance filters are the shared component. Under our MAX hypothesis, the shared component is the whole FRF circuit—everything highlighted in gray in Figure 1. It is shared in the sense that it can process both luminance- and contrast-modulated stimuli, but it is not engaged in all experimental groups at all times. This is one important

difference between the MAX hypothesis and the two earlier proposals. Specifically, the FRF pathway is not engaged during LM training in Group 1. Our working hypothesis is that when the two pathways are equally active, the quasi-linear pathway is selected by default. The support for this hypothesis is twofold. First, there is evidence that the quasi-linear pathway processes LM information faster than the FRF pathway (Derrington, Badcock, & Henning, 1993). This temporal advantage may break the tie at the integration stage (Wilson et al., 1992). Second, Edward and Badcock's (1995) finding that CM dots seem not to interfere with LM dots of calibrated contrast also suggests that LM signals are processed by the quasi-linear pathway when the LM and CM stimuli are approximately equally difficult. For our experimental design, we hypothesize that such tie occurs at the beginning of training with LM stimuli in Group 1, but no tie occurs during the LM post-test in Group 2. In the latter case, the extensive CM training has strengthened the FRF pathway so that it now responds more strongly than the untrained quasi-linear pathway to our (calibrated) LM stimuli.

The third empirical constraint is that there appears to be *full* transfer of CM training to LM test. As we argued in the Introduction section, this seems problematic for both the hierarchical and the early-filter interpretations. To recapitulate, the carrier of the non-Fourier stimuli moves in an orthogonal direction (Zanker, 1999) or has higher spatial frequency (Chen et al., 2009) than the direction or frequency of the Fourier stimuli. Thus, the non-Fourier training affects units tuned for values that are not optimal for the subsequent Fourier test. This predicts a drop in performance followed by subsequent recovery. The MAX hypothesis, on the other hand, accounts for the full transfer in a straightforward manner. The same substrate— the FRF pathway—determines the behavioral thresholds both before and after the CM → LM switch. This predicts no drop in performance (that is, full transfer) and no subsequent learning. This is exactly what was found (Figure 5, middle). The extended post-test in our experiment played a key role in testing this prediction.

The fourth empirical constraint is that there appears to be *full* LM → CM specificity. Again, this seems problematic for both the hierarchical and the early-filter interpretations. Under both proposals, the shared components— the early motion extractors or filters—are engaged in all experimental conditions at all times. This predicts transfer in both directions: from CM to LM (which agrees with the data) and from LM to CM (which does not). By contrast, the MAX hypothesis gives a natural account of the full LM → CM specificity. The shared components—the FRF circuits—are not engaged during LM training, as discussed above.

This account depends on an auxiliary assumption: The MAX elements in Figure 1f are assumed to gate not only which activation values are propagated to the decision stage but also which circuits are changed on a given trial.

Our working hypothesis is that learning occurs in only one channel *per alignment group* per trial—the channel with maximal activation within the group. Multiple alignment groups can still learn in parallel. Thus, the training sessions in Group 1 change the quasi-linear channels only; the FRF channels remain in their pre-training state. This assumption is reminiscent of the well-documented attentional influences on perceptual learning (e.g., Ahissar & Hochstein, 1993, 2002; Gilbert et al., 2001; Paffen, Verstraten, & Vidnyanszky, 2008; Roelfsema, van Ooyen, & Watanabe, 2010; Vidnyanszky & Sohn, 2005) and on learning more generally (e.g., Kruschke, 2003; Mackintosh, 1975). It is not contradicted by the evidence of "task-irrelevant learning" with subliminal stimuli (Seitz & Watanabe, 2005; Watanabe et al., 2001) because the MAX hypothesis concerns the selection among the cue-specific channels aligned for the same direction of motion, whereas the experiments of Watanabe et al. (2001) involve multiple directions.

The lack of LM → CM transfer also suggests an early plasticity cite. A secondary goal of the present study was to evaluate the prediction that fine discrimination training may transfer better than coarse discrimination training. The current data do not support this prediction. The transfer index in Group 1 is not significantly different from zero. This replicates the complete lack of transfer in the two coarse discrimination studies (Chen et al., 2009; Zanker, 1999). Taken at face value, this suggests that LM training induces no changes at the late, form-cue invariant stages in Figure 1 regardless of the task. This is indeed a very natural interpretation. However, the substantial individual differences in the transfer index opens the possibility that late plasticity occurs in some observers but not others.[9] Moreover, we found significant transfer in a related experiment (Hayes & Petrov, 2009). As we argued earlier, it is quite possible that the population mean for the LM → CM transfer index is between 15% and 45%. Note that this would be entirely consistent with the MAX hypothesis, or rather with a combination of MAX and selective reweighting (Dosher & Lu, 1998; Petrov et al., 2005). According to this combined proposal, there are two separate plasticity cites—one before and one after the MAX integration stage. The early plasticity produces massive transfer from CM to LM and lack of post-test learning, whereas the late plasticity produces modest transfer in both directions.

There are two ways in which perceptual learning can transfer from one stimulus class to another. Direct transfer occurs when the d' immediately after the switch is higher than the initial d'. Indirect transfer occurs when learning is faster during the post-test than the first training period (Liu & Weinshall, 2000). In this article, we focused on direct transfer as quantified by the transfer index in Equation 6. As we saw, it is not statistically significant in Group 1. However, there seems to be indirect transfer as indicated by the steep post-test learning curve (Figure 5, top, trials 3745–4992). The indirect transfer can be quantified by the ratio of the training and post-test time constants τ_1/τ_2 in Equation 5. This ratio has a median of 11 in our bootstrap samples and is greater than 5.5 in 95% of them. This indirect transfer suggests a late plasticity site (Jacobs, 2009; Liu & Weinshall, 2000). This is a topic for future research.

Related research

Asymmetric patterns of transfer of perceptual learning have been reported in monocular and binocular motion systems (Lu, Chu, Dosher, & Lee, 2005) and in orientation discrimination in clear and noisy displays (Dosher & Lu, 2005). A related study (Dosher & Lu, 2006) reports negligible learning effects in a discrimination task with luminance-defined letters but significant learning effects with texture-defined letters. Asymmetric effects on first- and second-order motion can also be induced by attentional manipulations (e.g., Allen & Ledgeway, 2003; Lu, Liu, & Dosher, 2000; Whitney & Bressler, 2007) and adaptation (e.g., Ashida et al., 2007; Whitney & Bressler, 2007). In all these cases, the asymmetry has been interpreted as evidence for dissociable pathways and/or learning mechanisms.

The interaction between first- and second-order visual processing has been investigated in a variety of psychophysical studies of masking (e.g., Edwards & Badcock, 1995), subthreshold facilitation (e.g., Lu & Sperling, 1995b), adaptation (e.g., Ledgeway & Smith, 1997; Nishida, Ledgeway et al., 1997; Turano, 1991), induced motion (e.g., Nishida, Edwards, & Sato, 1997), motion aftereffect (e.g., Nishida, Ashida, & Sato, 1994; Nishida & Sato, 1995; Schofield et al., 2007), plaids (e.g., Stoner & Albright, 1992; Victor & Conte, 1992), as well as static textures (e.g., Kingdom, Prins, & Hayes, 2003; Schofield & Georgeson, 1999; Schofield & Yates, 2005). We focus our discussion to those studies that are most relevant and/or contradictory to our results.

Edwards and Badcock (1995) measured coarse direction discrimination thresholds with random dot kinematograms consisting of various mixtures of LM (solid gray) dots and CM (checkerboard) dots. The LM dot contrast relative to the static noise background was calibrated to equalize the coherence thresholds of pure LM and pure CM stimuli. An asymmetric masking effect was found: The addition of incoherently moving LM dots elevated the CM threshold, whereas incoherent CM dots had no effect on the LM threshold. This finding—LM masks CM but not vice versa—may appear to be the exact opposite of our asymmetric transfer of learning. Upon closer examination, however, Edwards and Badcock's (1995) data seem entirely consistent with our MAX hypothesis and the dual-pathway architecture in Figure 1. The filter–rectify–filter channels are affected by both luminance and contrast modulations. This requires that CM signals must be affected by LM noise. On the other hand, the quasi-linear

pathway is blind to contrast modulations. This allows that LM signals can be unaffected by CM noise, *but only when the LM thresholds are determined by the quasi-linear pathway*. The latter condition depends on the MAX integration stage. When the stimuli are calibrated so that the two pathways are approximately equally active, we hypothesize that ties are resolved in favor of the quasi-linear pathway, as discussed above. It seems likely that the quasi-linear pathway determined the LM threshold in Edwards and Badcock's (1995) study. This is analogous to the training phase in our Group 1. On this interpretation, the LM threshold was not affected by CM noise (Edwards & Badcock, 1995) for the same reason that our LM training did not transfer to CM post-test. The situation reverses in Group 2 and this explains the apparent contradiction with Edwards and Badcock's (1995) data. To reiterate, we hypothesize that the LM post-test in Group 2 reflects the operation of the (strengthened) FRF circuits. Transported to the masking framework, the MAX hypothesis makes two predictions. First, if the LM contrast is calibrated as in Edwards and Badcock's (1995) study and then the observers practice with CM stimuli for several days, CM noise is predicted to mask the LM signal. Second, the same masking is predicted to occur even without CM training when the contrast of the LM dots is significantly lower than the contrast needed to achieve parity with the CM dots.

Nishida, Ledgeway et al. (1997) report an asymmetric pattern of cross-adaptation with LM and CM motion. Adaptation to CM motion had no significant effect on the subsequent LM detection thresholds, whereas adaptation to LM motion sometimes raised the CM detection thresholds. This asymmetry is consistent with the dual-pathway architecture if one assumes that the LM detection thresholds are determined by the quasi-linear pathway and that adaptation affects the cue-dependent circuits prior to the MAX integration stage (Shin'ya Nishida, personal communication, October 25, 2010). In agreement with the latter assumption, there is evidence that motion adaptation can occur even when the adapted motion is not visible (e.g., Maruya, Watanabe, & Watanabe, 2008; Nishida & Sato, 1992). However, the LM → CM cross-adaptation effects were "very weak and/or non-spatial-frequency selective" (Nishida, Ledgeway et al., 1997, p. 2692). The lack of spatial-frequency selectivity suggests that adaptation may also affect the form-cue invariant stages (Figure 1f). This is consistent with the symmetric cross-adaptation effects found in other studies (Ledgeway & Smith, 1997; Turano, 1991).

Schofield et al. (2007) report an asymmetric pattern of transfer of the dynamic motion aftereffect (dMAE[10]). The adaptation to an LM inducer transferred partially (48%) to a flickering CM test grating, whereas CM adaptation transferred weakly (28%) to LM (Schofield et al., 2007, Table 1, averaged across their 3 observers). This asymmetry is at odds with the asymmetric transfer of perceptual learning in our study (and those of Chen et al., 2009;

Zanker, 1999). It is also at odds with the findings of both Lu, Sperling, and Beck (1997), who found no dMAE transfer, and Nishida and Sato (1995), who found that higher order cues can impose a dMAE on LM tests. Furthermore, symmetric transfer is reported for the related tilt and contrast-reduction aftereffects (Cruickshank & Schofield, 2005; Georgeson & Schofield, 2002), but "we should not necessarily expect all aftereffects to follow the same pattern of transfer" (Schofield et al., 2007, p. 9). More research in this area is clearly necessary. One promising idea is that the MAE may arise from gain-control adaptation of broadly tuned inhibitory interneurons (Grunewald & Lankheet, 1996).

Conclusion

We obtained a strongly asymmetric pattern of transfer of perceptual leaning of luminance- and contrast-modulated motion in a fine direction discrimination task. CM training seemed to transfer fully to the LM post-test, but there was no significant transfer from LM to CM. The pattern of post-switch learning during the extended post-test was asymmetric as well. These strong asymmetries suggest a dual-pathway architecture with Fourier channels sensitive only to LM signals and non-Fourier channels sensitive to both LM and CM. We hypothesize that the channels tuned for the same motion direction, but different carriers are integrated using a MAX operation.

Appendix A

Neurophysiological evidence

This appendix evaluates various aspects of the proposed dual-pathway architecture with respect to single-cell recordings in cortical neurons in monkeys (Albright, 1992; Churan & Ilg, 2001; O'Keefe & Movshon, 1998) and cats (e.g., Mareschal & Baker, 1999; Zhou & Baker, 1993; see Baker & Mareschal, 2001, for review). All these studies report that a substantial fraction of the neurons in motion-sensitive areas seems directionally selective to LM but not CM stimuli. These neurons can be labeled as "pure first" and they are the likely substrate of the quasi-linear pathway in Figure 1.

All theoretical developments in this article depend on the assumption that the filter–rectify–filter (FRF) pathway is sensitive to both LM and CM stimuli. In addition to the psychophysical evidence discussed in the main text (e.g., Edwards & Badcock, 1995), this assumption rests on two complementary neurophysiological findings: the relative abundance of "mixed" neurons and the relative scarcity of "pure second" ones. With respect to the former, all studies

cited above report a large population of cells that seem directionally selective to both Fourier and non-Fourier stimuli. Specifically, if we pool the three monkey samples (Albright, 1992; Churan & Ilg, 2001; O'Keefe & Movshon, 1998), the share of these "mixed" neurons is 9% in area V1 (3/34), 44% (156/356) in MT, and 50% (34/68) in MST. Note that these proportions increase along the cortical hierarchy. As to "pure second" neurons, only 3 out of 207 MT neurons in the sample of O'Keefe and Movshon (1998) and 9 out of 106 MT and MST neurons in that of Churan and Ilg (2001) seemed to respond selectively to second- but not first-order motion. Moreover, the direction indices of these neurons were "rather low" (Churan & Ilg, 2001, p. 2302) and their responses to second-order stimuli were too close to the background firing rate to be behaviorally relevant (O'Keefe & Movshon, 1998).

The scarcity of "pure second" neurons should be interpreted with caution because there may be a sampling bias in favor of cells that respond vigorously to LM gratings and bars of light—both of which are standard search stimuli in single-cell recordings. Baker and Mareschal (2001) explicitly acknowledge that all neurons in their sample were always first characterized with sinewave gratings and thus necessarily responded to LM patterns. This is why we did not include the cat data in the counts above. Churan and Ilg (2001) and O'Keefe and Movshon (1998) are less explicit about their methods, but the fact that both studies found at least some "pure second" neurons suggests that they avoided this problem. Still, the extent to which the samples reported in the literature are representative of the relevant neuronal populations in the brain remains an open question.

These methodological difficulties notwithstanding, there are theoretical reasons to expect that "pure second" neurons are relatively rare. It is hard to achieve exact cancellation of all Fourier motion energy at all directions and frequencies.[11] The "purity" of a non-Fourier detector depends on the opacity of the FRF cascade to luminance modulations. The computer simulations (Edwards & Badcock, 1995; Wilson et al., 1992) cited in the Introduction section indicated that an FRF cascade with orthogonal filters is not opaque to LM dots and plaids. If the preferred orientations of the early and late filters were not orthogonal, the transfer functions of the respective filters would overlap more, making it easier for LM signals to pass. The neurophysiological evidence suggests a substantial overlap. Mareschal and Baker (1999) measured the responses to contrast-modulated motion stimuli whose carrier and envelope orientations were manipulated independently. They found no significant correlation between the preferred carrier orientation and the preferred envelope orientation. In 9 of 22 neurons, the two preferences differed by less than 30 degrees (Mareschal & Baker, 1999, Figure 9C). All neurons in this sample were screened to have at least 3-octave separation between the carrier and envelope spatial frequencies. When this constraint is relaxed, cortical neurons have been found to respond to envelope stimuli with closer envelope and carrier frequencies (O'Keefe & Movshon, 1998).

Finally, some of the "mixed" neurons may be receiving inputs from earlier "pure first" neurons. This is the likely substrate for the form-cue invariant units at the integration stage (Figure 1f; see also Albright, 1992; Zhou & Baker, 1993). This is consistent with the aforementioned finding that higher cortical areas contain apparently higher proportions of "mixed" cells. Note also that under our MAX integration hypothesis, form-cue invariance does not imply that the firing rates are completely independent of the cue. The MAX units in Figure 1f are form-cue invariant in a weaker sense. Namely, they are sharply tuned for the direction of motion and broadly tuned for the cue attributes. Some residual form-cue dependence remains, and the asymmetric transfer of learning suggests that it can be modified with practice. Even critics of form-cue invariance concede that "a subpopulation of MT neurons may be, in a broad sense, form-cue invariant, even if this property represents the exception rather than the rule" (O'Keefe & Movshon, 1998, p. 316). This is precisely what we expect in the architecture outlined in Figure 1—many more neurons are cue specific than cue invariant, because cue invariance is achieved by combining the outputs of multiple cue-specific neurons.

Appendix B

Calibration

This appendix addresses the concern that our data might be contaminated by first-order artifacts in our second-order stimuli. To be completely invisible to Fourier mechanisms, the second-order stimuli must be perfectly isoluminant. The field has developed sophisticated methods (e.g., Anstis & Cavanagh, 1983; Lu & Sperling, 2001a) to calibrate the stimuli as close as practically possible to this theoretical ideal. These fall into two categories: instrumentation calibration and observer-dependent calibration. We fully acknowledge the importance of the former, and the Methods section describes our efforts in this regard. However, we have conceptual reservations about the latter. The stated purpose of observer-dependent calibration (e.g., Lu & Sperling, 2001a) is to correct for non-linearities in the early visual system itself. For example, the transduction process in the photoreceptors in the retina is not perfectly linear (e.g., MacLeod, Williams, & Makous, 1992), and this induces first-order artifacts (e.g., Scott-Samuel & Georgeson, 1999; Smith & Ledgeway, 1997). It is argued that such distortions are not part of the motion-processing system per se and, therefore, should be corrected. The problem is that this puts us on a slippery slope. Where does one draw

the line? Is the processing in the thalamus relevant, for instance? Another problem is that observer-dependent calibration introduces what philosophers of science (Hanson, 1958) refer to as *theory-ladenness of observation*. An adaptive procedure adds titrated components to the images until the discrimination performance drops to chance level or otherwise satisfies some pre-defined criterion (e.g., Lu & Sperling, 2001a). It seems circular to argue on the basis of behavioral data obtained with such calibrated stimuli that there are two *independent* motion-processing pathways. Rather, the independence is assumed a priori and any interactions are treated as experimental imperfections. A third problem is that the sensitive calibration methods (e.g., Lu & Sperling, 2001a) apply to gratings only, and thereby restrict the empirical base unnecessarily. This is because they depend on cancellation of modulations at different phases, which in turn depends on the symmetry of the sinusoidal gratings. We are aware of no cancellation method that can be applied to the stochastic textures in our stimuli. This prevented us from measuring the magnitude of the "impurities" in the Fourier channel explicitly. It seems highly unlikely, however, that our CM stimuli "leak" into the first-order channel because if they did, LM training would transfer to CM, contrary to the data from Group 1. Furthermore, dynamic noise carriers have been shown to reduce the artifacts in some cases (Smith & Ledgeway, 1997).

Appendix C

Hierarchical non-linear regression

Here we describe the hierarchical non-linear regression that led to Table 1. We also describe the bootstrap procedure that generated the confidence intervals and standard deviation estimates throughout the text.

All analyses were performed on group-average data. The full data set in Figure 5 consists of 96 points in 4 profiles (2 groups by 2 difficulty levels). Equation 5 describes one such profile with 8 parameters. Thus, a fully saturated model has 32 parameters ($R^2 = 0.922$, RMSE = 0.089). Six parameters were eliminated by using the same time constants τ_1 and τ_2 for all 4 profiles ($R^2 = 0.916$, RMSE = 0.092). In addition, the d' for the easy discrimination ($\Delta = 11°$) seems proportional to the d' for the difficulty discrimination ($\Delta = 8°$). Thus, a single proportionality coefficient k replaced the 12 parameters for the easy d' levels ($R^2 = 0.914$, RMSE = 0.093). Three additional degrees of freedom were eliminated by sharing d_1, d_2, and d_3 across the two groups ($R^2 = 0.914$, RMSE = 0.093; the group-specific estimates are $d_{1,1} = 0.76$, $d_{2,1} = 0.98$, $d_{3,1} = 0.93$; $d_{1,2} = 0.79$, $d_{2,2} = 1.03$, and $d_{3,2} = 0.93$, where the second subscript denotes the group). Finally, the lack of learning in the post-test segment in Group 2 eliminated

$d_{24,2}$ ($R^2 = 0.913$, RMSE = 0.094). The small drop in R^2 caused by these reductions was amply justified by the gain in parsimony ($F < 1$ in all cases, n.s.). The resulting model has 11 free parameters listed in Table 1.

We used the bootstrap method (Efron & Tibshirani, 1993) to estimate the variance of the parameter estimates. To generate one bootstrap set, the participants were sampled with replacement, making sure that the set contained 10 individuals from each group. The 11-parameter regression model was then fitted to the group-averaged bootstrap data and various transfer indices were calculated (cf., Table 2). This procedure was repeated 1000 times to produce the confidence intervals and z-tests reported in the text.

Footnotes

[1]The specific proposals include full-wave rectification (e.g., Wilson, 1999), half-wave rectification (e.g., Solomon & Sperling, 1994), and half-squaring (e.g., Heeger, 1992a).

[2]More precisely, a softened variant in which several units remain active (k-WTA) is needed to model transparency (Kim & Wilson, 1994) and to increase the precision of the representation (Wilson et al., 1992).

[3]Within the precision of the figures of Chen et al. (2009). The article is in Chinese and we are not sure of the details because we cannot read the text. Our presentation is based on the English abstract and figure captions, as well as a brief email exchange with Yifeng Zhou.

[4]Our experiment was designed (Hayes & Petrov, 2009) and the data were collected before we become aware of the study of Chen et al. (2009). The latter also uses CM stimuli but has no extended post-test and thus does not measure the amount of post-switch learning. A range of temporal frequencies are surveyed instead. In this and other respects, the experimental design of Chen et al. is complementary to ours.

[5]Specifically, we fixed $\beta = 0.90$, ran three observers on a sequence of blocks covering a range of α values, and

estimated the value (0.13) for which the average first-order d' matched the average second-order d'.

[6]Less than 0.1% of the total number of trials were repeated, and most of them came from two individual observers.

[7]We also calculated an alternative transfer index based on between-group comparisons within a given stimulus class. Both indices pointed to the same qualitative conclusions.

[8]In the group that trained with LM motion.

[9]Part of these differences may also reflect other factors such as attention and strategy.

[10]Static motion aftereffect (sMAE) is the phenomenon that, following prolonged viewing of a moving inducer, a static test pattern appears to move in the opposite direction. The dynamic MAE is the analogous phenomenon with a flickering test pattern. LM adaptation induces both types of MAE, whereas CM adaptation induces only dMAE (Nishida & Sato, 1995).

[11]The careful calibration necessary to minimize first-order artifacts in second-order stimuli is another manifestation of this difficulty (see Appendix B).

References

Adelson, E. H., & Bergen, J. R. (1985). Spatiotemporal energy models for the perception of motion. *Journal of the Optical Society of America A, 2,* 284–299.

Adelson, E. H., & Movshon, J. A. (1982). Phenomenal coherence of moving visual patterns. *Nature, 300,* 523–535.

Ahissar, M., & Hochstein, S. (1993). Attentional control in early perceptual learning. *Proceedings of the National Academy of Sciences of the United States of America, 90,* 5718–5722.

Ahissar, M., & Hochstein, S. (1996). Learning pop-out detection: Specificities to stimulus characteristics. *Vision Research, 36,* 3487–3500.

Ahissar, M., & Hochstein, S. (1997). Task difficulty and the specificity of perceptual learning. *Nature, 387,* 401–406.

Ahissar, M., & Hochstein, S. (2002). The role of attention in learning simple visual tasks. In M. Fahle & T. Poggio (Eds.), *Perceptual learning* (pp. 253–272). Cambridge, MA: MIT Press.

Ahissar, M., & Hochstein, S. (2004). The reverse hierarchy theory of visual perceptual learning. *Trends in Cognitive Sciences, 8,* 457–464.

Albright, T. D. (1992). Form-cue invariant motion processing in primate visual cortex. *Science, 255,* 1141–1143.

Allen, H. A., & Ledgeway, T. (2003). Attentional modulation of threshold sensitivity to first-order motion and second-order motion patterns. *Vision Research, 43,* 2927–2936.

Anstis, S., & Cavanagh, P. (1983). A minimum motion technique for judging equiluminance. In J. D. Mollon & L. T. Sharpe (Eds.), *Colour vision: Physiology and psychophysics* (pp. 155–166). London: Academic Press.

Ashida, H., Lingnau, A., Wall, M. B., & Smith, A. T. (2007). fMRI adaptation reveals separate mechanisms for first-order and second-order motion. *Journal of Neurophysiology, 97,* 1319–1325.

Baker, C. L., & Mareschal, I. (2001). Processing of second-order motion stimuli in the visual cortex. *Progress in Brain Research, 134,* 171–191.

Ball, K., & Sekuler, R. (1987). Direction-specific improvement in motion discrimination. *Vision Research, 27,* 953–965.

Baloch, A. A., Grossberg, S., Mingolla, E., & Nogueira, C. A. M. (1999). Neural model of first-order and second-order motion perception and magnocellular dynamics. *Journal of the Optical Society of America A, 16,* 953–978.

Benton, C. P. (2002). Gradient-based analysis of non-Fourier motion. *Vision Research, 42,* 2869–2877.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision, 10,* 433–436.

Britten, K. H., Shadlen, M. N., Newsome, W. T., & Movshon, J. A. (1992). The analysis of visual motion: A comparison of neuronal and psychophysical performance. *Journal of Neuroscience, 12,* 4745–4765.

Cameron, E. L., Baker, C. L., & Boulton, J. C. (1992). Spatial frequency selective mechanism underlying the motion aftereffect. *Vision Research, 32,* 561–568.

Carandini, M., Heeger, D. J., & Movshon, J. A. (1999). Linearity and gain control in V1 simple cells. In P. S. Ulinski, E. G. Jones, & A. Peters (Eds.), *Cerebral cortex: Vol. 13. Models of cortical circuits* (pp. 445–447). New York: Kluwer Academic/Plenum Publishers.

Cavanagh, P. (1992). Attention-based motion perception. *Science, 257,* 1563–1565.

Cavanagh, P., Arguin, M., & von Grüunau, M. W. (1989). Interattribute apparent motion. *Vision Research, 29,* 1197–1204.

Cavanagh, P., & Mather, G. (1989). Motion: The long and short of it. *Spatial Vision, 4,* 103–129.

Chen, R., Qiu, Z.-P., Zhang, Y., & Zhou, Y.-F. (2009). Perceptual learning and transfer study of first- and second-order motion direction discrimination (in Chinese). *Progress in Biochemistry and Biophysics, 36,* 1442–1450.

Chubb, C., & Landy, M. S. (1991). Orthogonal distribution analysis: A new approach to the study of texture

perception. In M. S. Landy & J. A. Movshon (Eds.), *Computational models of visual processing* (pp. 291–301). Cambridge, MA: MIT Press.

Chubb, C., & Sperling, G. (1988). Drift-balanced random dot stimuli: A general basis for studying non-Fourier motion perception. *Journal of the Optical Society of America A, 5,* 1986–2007.

Churan, J., & Ilg, U. J. (2001). Processing of second-order motion stimuli in primate middle temporal area and medial superior temporal area. *Journal of the Optical Society of America A, 18,* 2297–2306.

Colombo, E., & Derrington, A. (2001). Visual calibration of CRT monitors. *Displays, 22,* 87–95.

Crist, R. E., Kapadia, M. K., Westheimer, G., & Gilbert, C. D. (1997). Perceptual learning of spatial location: Specificity for orientation, position, and context. *The Journal of Physiology, 78,* 2889–2894.

Cruickshank, A. G., & Schofield, A. J. (2005). Transfer of tilt after-effects between second-order cues. *Spatial Vision, 18,* 379–397.

Derrington, A. M., Badcock, D. R., & Henning, G. B. (1993). Discriminating the direction of second-order motion at short stimulus durations. *Vision Research, 33,* 1785–1794.

De Valois, R. L., & De Valois, K. K. (1988). *Spatial vision.* New York: Oxford University Press.

Dosher, B. A., & Lu, Z.-L. (1998). Perceptual learning reflects external noise filtering and internal noise reduction through channel reweighting. *Proceedings of the National Academy of Sciences of the United States of America, 95,* 13988–13993.

Dosher, B. A., & Lu, Z.-L. (2005). Perceptual learning in clear displays optimizes perceptual expertise: Learning the limiting process. *Proceedings of the National Academy of Sciences of the United States of America, 102,* 5286–5290.

Dosher, B. A., & Lu, Z.-L. (2006). Level and mechanisms of perceptual learning: Learning first-order luminance and second-order texture objects. *Vision Research, 46,* 1996–2007.

Edwards, M., & Badcock, D. R. (1995). Global motion perception: No interaction between the first- and second-order motion pathways. *Vision Research, 35,* 2589–2602.

Efron, B., & Tibshirani, R. J. (1993). *An introduction to the bootstrap.* New York: Chapman and Hall.

Fahle, M., & Edelman, S. (1993). Long-term learning in Vernier acuity: Effects of stimulus orientation, range, and feedback. *Vision Research, 33,* 397–412.

Fahle, M., & Poggio, T. (Eds.) (2002). *Perceptual learning.* Cambridge, MA: MIT Press.

Fine, I., & Jacobs, R. A. (2002). Comparing perceptual learning across tasks: A review. *Journal of Vision,* 2(2):5, 190–203, http://www.journalofvision.org/content/2/2/5, doi:10.1167/2.2.5. [PubMed] [Article]

Gawne, T. J., & Martin, J. M. (2002). Responses of primate visual cortical V4 neurons to simultaneously presented stimuli. *Journal of Neurophysiology, 88,* 1128–1135.

Georgeson, M. A., & Schofield, A. J. (2002). Shading and texture: Separate information channels with a common adaptation mechanism? *Spatial Vision, 16,* 59–76.

Gilbert, C. D., Sigman, M., & Crist, R. E. (2001). The neural basis of perceptual learning. *Neuron, 31,* 681–697.

Gold, J. M., Bennett, P. J., & Sekuler, A. B. (1999). Signal but not noise changes with perceptual learning. *Nature, 402,* 176–178.

Graham, N., Beck, J., & Sutter, A. (1992). Nonlinear processes in spatial-frequency channel models of perceived texture segregation: Effects of sign and amount of contrast. *Vision Research, 32,* 719–743.

Grunewald, A., & Lankheet, M. J. M. (1996). Orthogonal motion after-effect illusion predicted by a model of cortical motion processing. *Nature, 384,* 358–360.

Grzywacz, N. M., Watamaniuk, S. N. J., & McKee, S. P. (1995). Temporal coherence theory for the detection and measurement of visual motion. *Vision Research, 35,* 3183–3203.

Hanson, N. R. (1958). *Patterns of discovery: An inquiry into the conceptual foundations of science.* Cambridge, UK: Cambridge University Press.

Hayes, T., & Petrov, A. A. (2009). Perceptual learning transfers from luminance- to contrast-defined motion [Abstract]. *Journal of Vision, 9*(8):884, 884a, http://www.journalofvision.org/content/9/8/884, doi:10.1167/9.8.884.

Heathcote, A. J., Brown, S. D., & Mewhort, D. J. K. (2000). The power law repealed: The case for an exponential law of practice. *Psychonomic Bulletin & Review, 7,* 185–207.

Heeger, D. J. (1992a). Half-squaring in responses of cat striate cells. *Visual Neuroscience, 9,* 427–443.

Heeger, D. J. (1992b). Normalization of cell responses in cat striate cortex. *Visual Neuroscience, 9,* 181–197.

Hol, K., & Treue, S. (2001). Different populations of neurons contribute to the detection and discrimination of visual motion. *Vision Research, 41,* 685–834.

Huang, X., Lu, H., Tjan, B. S., Zhou, Y.-F., & Liu, Z. (2008). Motion perceptual learning: When only task-relevant information is learned. *Journal of Vision, 7*(10):14, 1–10, http://www.journalofvision.org/content/7/10/14, doi:10.1167/7.10.14. [PubMed] [Article]

Ilan, L., Ferster, D., Poggio, T., & Riesenhuber, M. (2004). Intracellular measurements of spatial integration and

the MAX operation in complex cells of the cat primary visual cortex. *Journal of Neurophysiology, 92,* 2704–2713.

Jacobs, R. A. (2009). Adaptive precision pooling of model neuron activities predicts the efficiency of human visual learning. *Journal of Vision, 9*(4):22, 1–15, http://www.journalofvision.org/content/9/4/22, doi:10.1167/9.4.22. [PubMed] [Article]

Jazayeri, M., & Movshon, J. A. (2006). Optimal representation of sensory information by neural populations. *Nature Neuroscience, 9,* 690–696.

Johnston, A., & Clifford, C. W. G. (1995). A unified account of three apparent motion illusions. *Vision Research, 35,* 1109–1123.

Johnston, A., McOwan, P. W., & Buxton, H. (1992). A computational model of the analysis of some first-order and second-order motion patterns by simple and complex cells. *Proceedings of the Royal Society of London B, 250,* 297–306.

Karni, A., & Sagi, D. (1991). Where practice makes perfect in texture discrimination: Evidence for primary visual cortex plasticity. *Proceedings of the National Academy of Sciences of the United States of America, 88,* 4966–4970.

Kim, J., & Wilson, H. R. (1994). A model for motion coherence and transparency. *Visual Neuroscience, 11,* 1205–1220.

Kingdom, F. A. A., Prins, N., & Hayes, A. (2003). Mechanism independence for texture-modulation detection is consistent with a filter–rectify–filter mechanism. *Visual Neuroscience, 20,* 65–76.

Klein, S. A., Hu, Q. J., & Carney, T. (1996). The adjacent pixel nonlinearity: Problems and solutions. *Vision Research, 36,* 3167–3181.

Kruschke, J. K. (2003). Attention in learning. *Current Directions in Psychological Science, 12,* 171–175.

Law, C.-T., & Gold, J. I. (2008). Neural correlates of perceptual learning in a sensory motor, but not a sensory, cortical area. *Nature Neuroscience, 11,* 505–513.

Law, C.-T., & Gold, J. I. (2009). Reinforcement learning can account for associative and perceptual learning on a visual-decision task. *Nature Neuroscience, 12,* 655–663.

Ledgeway, T., & Smith, A. T. (1997). Changes in perceived speed following adaptation to first-order and second-order motion. *Vision Research, 37,* 215–224.

Liu, Z. (1999). Perceptual learning in motion discrimination that generalizes across motion directions. *Proceedings of the National Academy of Sciences of the United States of America, 96,* 14085–14087.

Liu, Z., & Weinshall, D. (2000). Mechanisms of generalization in perceptual learning. *Vision Research, 40,* 97–109.

Lu, Z.-L., Chu, W., Dosher, B. A., & Lee, S. (2005). Independent perceptual learning in monocular and binocular motion systems. *Proceedings of the National Academy of Sciences of the United States of America, 102,* 5624–5629.

Lu, Z.-L., & Dosher, B. A. (2008). Characterizing observers using external noise and observer models: Assessing internal representations with external noise. *Psychological Review, 115,* 44–82.

Lu, Z.-L., Liu, C. Q., & Dosher, B. A. (2000). Attention mechanisms for multi-location first- and second-order motion perception. *Vision Research, 40,* 173–176.

Lu, Z.-L., Liu, J., & Dosher, B. A. (2010). Modeling mechanisms of perceptual learning with augmented Hebbian re-weighting. *Vision Research, 50,* 375–390.

Lu, Z.-L., & Sperling, G. (1995a). Attention-generated apparent motion. *Nature, 377,* 237–239.

Lu, Z.-L., & Sperling, G. (1995b). The functional architecture of human visual motion perception. *Vision Research, 35,* 2697–2722.

Lu, Z.-L., & Sperling, G. (2001a). Sensitive calibration and measurement procedures based on the amplification principle in motion perception. *Vision Research, 41,* 2355–2374.

Lu, Z.-L., & Sperling, G. (2001b). Three-systems theory of human visual motion perception: Review and update. *Journal of the Optical Society of America A, 18,* 2331–2370.

Lu, Z.-L., Sperling, G., & Beck, J. R. (1997). Selective adaptation of three motion systems. *Investigative Ophthalmology & Visual Science, 38,* 1105.

Mackintosh, N. J. (1975). A theory of attention: Variation in the associability of stimuli with reinforcement. *Psychological Review, 82,* 276–298.

MacLeod, D. I. A., Williams, D. R., & Makous, W. (1992). A visual non-linearity fed by single cones. *Vision Research, 32,* 347–363.

Mareschal, I., & Baker, C. L. (1999). Cortical processing of second-order motion. *Visual Neuroscience, 16,* 527–540.

Maruya, K., & Nishida, S. (2010). Feature invariant spatial pooling of first- and second-order motion signals for solution of aperture problem [Abstract]. *Journal of Vision, 10*(7):836, 836a, http://www.journalofvision.org/content/10/7/836, doi:10.1167/10.7.836.

Maruya, K., Watanabe, H., & Watanabe, M. (2008). Adaptation to invisible motion results in low-level but not high-level aftereffects. *Journal of Vision, 8*(11):7,

1–11, http://www.journalofvision.org/content/8/11/7, doi:10.1167/8.11.7. [PubMed] [Article]

Matthews, N., & Welch, L. (1997). Velocity-dependent improvements in single-dot direction discrimination. *Perception & Psychophysics, 59,* 60–72.

Nishida, S., Ashida, H., & Sato, T. (1994). Complete interocular transfer of motion aftereffect with flickering test. *Vision Research, 34,* 2707–2716.

Nishida, S., Edwards, M., & Sato, T. (1997). Simultaneous motion contrast across space: Involvement of second-order motion? *Vision Research, 37,* 199–214.

Nishida, S., Ledgeway, T., & Edwards, M. (1997). Dual multiple-scale processing for motion in the human visual system. *Vision Research, 37,* 2685–2698.

Nishida, S., & Sato, T. (1992). Positive motion after-effect induced by band-pass-filtered random-dot kinematograms. *Vision Research, 32,* 1635–1646.

Nishida, S., & Sato, T. (1995). Motion aftereffect with flickering test patterns reveals higher stages of motion processing. *Vision Research, 35,* 477–490.

Nowlan, S. J., & Sejnowski, T. J. (1995). A selection model for motion processing in area MT of primates. *Journal of Neuroscience, 15,* 1195–1214.

O'Keefe, L. P., & Movshon, A. J. (1998). Processing of first- and second-order motion signals by neurons in area mt of the macaque monkey. *Visual Neuroscience, 15,* 305–317.

Paffen, C. L. E., Verstraten, F. A. J., & Vidnyanszky, Z. (2008). Attention-based perceptual learning increases binocular rivalry suppression of irrelevant visual features. *Journal of Vision, 8*(4):25, 1–11, http://www.journalofvision.org/content/8/4/25, doi:10.1167/8.4.25. [PubMed] [Article]

Papathomas, T., Rosenthal, A. S., & Julesz, B. (2002). Neural models of motion perception. In G. K. Hung & K. J. Ciuffreda (Eds.), *Models of the visual system* (pp. 487–519). New York: Kluwer Academic/Plenum Publishers.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision, 10,* 437–442.

Petrov, A. A. (2009). Symmetry-based methodology for decision-rule identification in same–different experiments. *Psychonomic Bulletin & Review, 16,* 1011–1025.

Petrov, A. A., Dosher, B. A., & Lu, Z.-L. (2005). The dynamics of perceptual learning: An incremental reweighting model. *Psychological Review, 112,* 715–743.

Petrov, A. A., Dosher, B. A., & Lu, Z.-L. (2006). Perceptual learning without feedback in nonstationary contexts: Data and model. *Vision Research, 46,* 3177–3197.

Purushothaman, G., & Bradley, D. C. (2005). Neural population code for fine perceptual decisions in area MT. *Nature Neuroscience, 8,* 99–106.

Qian, N., Andersen, R. A., & Adelson, E. H. (1994). Transparent motion perception as detection of unbalanced motion signals: I. Psychophysics. *Journal of Neuroscience, 14,* 7357–7366.

Raiguel, S., Vogels, R., Mysore, S. G., & Orban, G. A. (2006). Learning to see the difference specifically alters the most informative V4 neurons. *Journal of Neuroscience, 26,* 6589–6602.

Regan, D., & Beverley, K. I. (1985). Postadaptation orientation discrimination. *Journal of the Optical Society of America A, 2,* 147–155.

Riesenhuber, M., & Poggio, T. (1999). Hierarchical models of object recognition in cortex. *Nature Neuroscience, 2,* 1019–1025.

Roelfsema, P. R., van Ooyen, A., & Watanabe, T. (2010). Perceptual learning rules based on reinforcers and attention. *Trends in Cognitive Sciences, 14,* 64–71.

Schofield, A. J., & Georgeson, M. A. (1999). Sensitivity to modulations of luminance and contrast in visual white noise: Separate mechanisms with similar behavior. *Vision Research, 39,* 2697–2716.

Schofield, A. J., Ledgeway, T., & Hutchinson, C. V. (2007). Asymmetric transfer of the dynamic motion aftereffect between first- and second-order cues and among different second-order cues. *Journal of Vision, 7*(8):1, 1–12, http://www.journalofvision.org/content/7/8/1, doi:10.1167/7.8.1. [PubMed] [Article]

Schofield, A. J., & Yates, T. A. (2005). Interactions between orientation and contrast modulations suggest limited cross-cue linkage. *Perception, 34,* 769–792.

Scott-Samuel, N. E., & Georgeson, M. A. (1999). Does early non-linearity account for second-order motion? *Vision Research, 39,* 2853–2865.

Seitz, A. R., & Watanabe, T. (2005). A unified model of perceptual learning. *Trends in Cognitive Science, 9,* 329–334.

Serre, Oliva, & Poggio (2007). A feedforward architecture accounts for rapid categorization. *Proceedings of the National Academy of Sciences of the United States of America, 104,* 6424–6429.

Seung, H. S., & Sompolinsky, H. (1993). Simple models for reading neuronal population codes. *Proceedings of the National Academy of Sciences of the United States of America, 90,* 10749–10753.

Shadlen, M. N., Britten, K. H., Newsome, W. T., & Movshon, J. A. (1996). A computational analysis of the relationship between neuronal and behavioral

responses to visual motion. *Journal of Neuroscience, 16,* 1486–1510.

Smith, A. T., & Ledgeway, T. (1997). Separate detection of moving luminance and contrast modulations: Fact or artifact? *Vision Research, 37,* 45–62.

Snowden, R. J., & Verstraten, F. A. J. (1999). Motion transparency: Making models of motion perception transparent. *Trends in Cognitive Sciences, 3,* 369–377.

Solomon, J. A., & Sperling, G. (1994). Full-wave and half-wave rectification in second-order motion perception. *Vision Research, 34,* 2239–2257.

Stoner, G. R., & Albright, T. D. (1992). Motion coherency rules are form-cue invariant. *Vision Research, 32,* 465–475.

Taub, E., Victor, J. D., & Conte, M. M. (1997). Nonlinear preprocessing in short-range motion. *Vision Research, 37,* 1459–1477.

The MathWorks (1999). *MATLAB user's guide (Computer software manual).* Natick, MA: The MathWorks.

Treue, S., & Maunsell, J. H. R. (1996). Attentional modulation of visual motion processing in cortical areas MT and MST. *Nature, 382,* 539–541.

Tsui, J. M. G., Hunter, J. N., Born, R. T., & Pack, C. C. (2010). The role of V1 surround suppression in MT motion integration. *Journal of Neurophysiology, 103,* 3123–3138.

Turano, K. (1991). Evidence for a common motion mechanism of luminance-modulated and contrast-modulated patterns: Selective adaptation. *Perception, 20,* 455–466.

Vaina, L. M., & Soloviev, S. (2004). First-order and second-order motion: Neurological evidence for neuroanatomically distinct systems. *Progress in Brain Research, 144,* 197–212.

Vaina, L. M., Sundareswaran, V., & Harris, J. G. (1995). Learning to ignore: Psychophysics and computational modeling of fast learning of direction in noisy motion stimuli. *Cognitive Brain Research, 2,* 155–163.

Van Horn, N., & Petrov, A. A. (2009). Perceptual learning of visual motion: The role of the spatial frequency of the carrier [Abstract]. *Journal of Vision, 9*(8):886, 886a, http://www.journalofvision.org/content/9/8/886, doi:10.1167/9.8.886.

van Santen, J. P. H., & Sperling, G. (1985). Elaborated Reichart detectors. *Journal of the Optical Society of America A, 2,* 300–320.

Victor, J. D., & Conte, M. M. (1992). Coherence and transparency of moving plaids composed of Fourier and non-Fourier gratings. *Perception & Psychophysics, 52,* 403–414.

Vidnyanszky, Z., & Sohn, W. (2005). Learning to suppress task-irrelevant visual stimuli with attention. *Vision Research, 45,* 677–685.

Watanabe, T., Náñez, J., & Sasaki, Y. (2001). Perceptual learning without perception. *Nature, 413,* 844–848.

Whitney, D., & Bressler, D. W. (2007). Second-order motion without awareness: Passive adaptation to second-order motion produces a motion aftereffect. *Vision Research, 47,* 569–579.

Wilson, H. R. (1999). Non-Fourier cortical processes in texture, form, and motion perception. In P. S. Ulinski, E. G. Jones, & A. Peters (Eds.), *Cerebral cortex: Vol. 13. Models of cortical circuits* (pp. 445–447). New York: Kluwer Academic/Plenum Publishers.

Wilson, H. R., Ferrera, V. P., & Yo, C. (1992). A psychophysically motivated model for two-dimensional motion perception. *Visual Neuroscience, 9,* 79–97.

Yu, A. J., Giese, M. A., & Poggio, T. A. (2002). Biophysiologically plausible implementations of the maximum operation. *Neural Computation, 14,* 2857–2881.

Yuille, A. L., & Grzywacz, N. M. (1989). A winner-take-all mechanism based on presynaptic inhibition feedback. *Neural Computation, 1,* 334–347.

Zanker, J. M. (1993). Theta motion: A paradoxical stimulus to explore higher order motion extraction. *Vision Research, 33,* 553–569.

Zanker, J. M. (1996). On the elementary mechanism underlying secondary motion processing. *Philosophical Transactions of the Royal Society of London B, 351,* 1725–1736.

Zanker, J. M. (1999). Perceptual learning in primary and secondary motion vision. *Vision Research, 39,* 1293–1304.

Zhou, Y.-X., & Baker, C. L. (1993). A processing stream in mammalian visual cortex neurons for non-Fourier responses. *Science, 261,* 98–101.