No second-order motion system sensitive to high temporal frequencies

Rémy Allard

Visual Psychophysics and Perception Laboratory,
Université de Montréal, Canada

Jocelyn Faubert

Visual Psychophysics and Perception Laboratory,
Université de Montréal, Canada

It has been shown that the perception of contrast-defined motion (i.e., a second-order stimulus) at high temporal frequencies cannot be explained solely by global distortion products (i.e., luminance artifacts due to preprocessing nonlinearities) processed by the first-order system. However, previous studies rejecting the first-order pathway hypothesis have assumed that the preprocessing nonlinearities are identical for all first-order motion units. If this is not the case, then introducing a nonlinearity within the stimulus could neutralize the *global* (i.e., mean) distortion product but would leave *residual* distortion products. We neutralized either global only or both global and residual distortion products by superimposing a luminance modulation onto the contrast modulation. At a temporal frequency too high for features to be tracked (15 Hz), we found a substantial texture (i.e., contrast-modulated) contribution to motion when neutralizing only global distortion products but not when neutralizing both global and residual distortion products. Furthermore, we found that the texture contribution to motion at this high temporal frequency, when it was not completely neutralized, depended on the phase difference between luminance and contrast modulations, which implied some common processing before the motion extraction stage. We concluded that the texture contribution to motion at high temporal frequencies was due to nonuniform preprocessing nonlinearities within the visual system, enabling first-order motion units to process distortion products, and not due to a dedicated second-order motion system.

Introduction

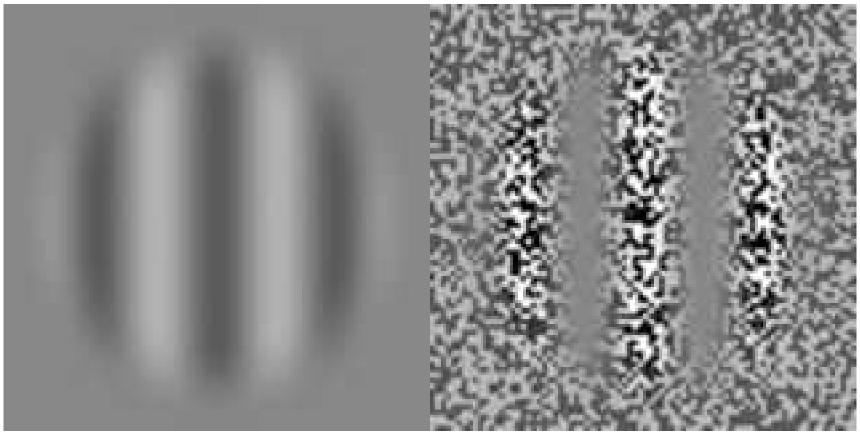
We have at least two motion systems based on fundamentally different computational strategies (Anstis & Mackay, 1980; Braddick, 1974; Braddick, Ruddock, Morgan, & Marr, 1980; Julesz, 1971). One directly processes spatiotemporal luminance modulations without requiring a preprocessing feature extraction stage (Adelson & Bergen, 1985; van Santen & Sperling, 1985; Watson & Ahumada, 1985). This motion system is often referred to as an energy-based, intensity-based, Fourier, linear, or first-order motion system and is typically found to be low-level, monocular, and fast (i.e., with a cutoff frequency near 12 Hz) (Lu & Sperling, 2001). The other motion system tracks the change in position of a feature that must first be extracted (Cavanagh, 1992). This motion system is usually referred to as a feature-tracking, third-order, high-level, or correspondence-based motion system and is known to be high-level, binocular, and slow (i.e., with a cutoff frequency near 3 Hz) (Lu & Sperling, 2001). In the present study, we will refer to these two motion systems as the first-order and feature-tracking motion systems, respectively.

Besides these two motion systems, many authors have suggested the existence of a second-order motion system (for reviews, see Burr & Thompson, 2011; Nishida, 2011) that would be sensitive to contrast-defined motion (Movie 1, right). However, the properties of this second-order motion system (assuming it exists) are still debated. For instance, some argue that it would be bandpass, i.e., sensitive to both low and high temporal frequencies (Lu & Sperling, 1995, 2001; Scott-Samuel & Georgeson, 1999); whereas others rather suggest that it would be lowpass, i.e., sensitive only to low temporal frequencies (Hutchinson & Ledgeway, 2006; Smith & Ledgeway, 1997).

There are massive amounts of evidence showing that luminance- and contrast-defined motion (Movie 1) are, under many conditions, processed by distinct motion systems. However, in many cases, they do not necessarily reveal properties of the second-order motion system because the perception of contrast-defined motion could be due to the feature-tracking

Citation: Allard, R., & Faubert, J. (2013) No second-order motion system sensitive to high temporal frequencies. *Journal of Vision, 13*(5):4, 1–14, http://www.journalofvision.org/content/13/5/4, doi:10.1167/13.5.4.

doi: 10.1167/13.5.4 Received July 26, 2012; published April 4, 2013 ISSN 1534-7362 © 2013 ARVO



Movie 1. Luminance- (left) and contrast-defined (right) motion.

motion system. For instance, contrary to the first-order motion system, it has been argued that the second-order motion system requires long stimulus durations (Derrington, Badcock, & Henning, 1993), cannot discriminate the direction of motion at detection threshold (Ledgeway & Hutchinson, 2005; Smith & Ledgeway, 1997), is temporally lowpass (Hutchinson & Ledgeway, 2006), does not induce optokinetic nystagmus (Harris & Smith, 1992), does not generate a static-motion aftereffect (Nishida & Sato, 1995), and is neuroanatomically distinct from the first-order system (Vaina & Soloviev, 2004). These properties differ from those of the first-order motion system and therefore must belong to another motion system. However, because they could belong to the feature-tracking motion system, they do not necessarily reveal properties of the second-order motion system.

Determining the properties of a dedicated second-order motion system is difficult given the existence of two other motion systems (i.e., first-order and feature-tracking) that are potentially responsible for some texture (e.g., contrast-modulation) contribution to motion. To simplify this matter, the present study focused on a temporal frequency that is too high for features to be tracked (15 Hz). At such a temporal frequency, the texture contribution to motion cannot be due to the feature-tracking motion system and must therefore be due to the first- or second-order motion system. The goal of the present study was to determine which of these motion systems enables the texture contribution to motion at a temporal frequency too high for features to be tracked.

Some studies suggest that texture contribution to motion at high temporal frequencies must be mediated by the first-order motion system either because there is no dedicated second-order motion system (Allard & Faubert, 2008a; Ashida, Seiffert, & Osaka, 2001; Derrington & Ukkonen, 1999; Ukkonen & Derrington, 2000) or because the second-order motion system is temporally lowpass (i.e., sensitive only to low temporal frequencies) (Hutchinson & Ledgeway, 2006; Smith & Ledgeway, 1997). Conversely, other studies suggest the existence of a dedicated second-order motion system that is bandpass and therefore could enable the texture contribution to motion at high temporal frequencies (e.g., Lu & Sperling, 1995, 2001; Scott-Samuel & Georgeson, 1999). Note that by using dynamic noise carriers Ledgeway and colleagues have found that the second-order motion system is temporally lowpass (Hutchinson & Ledgeway, 2006; Smith & Ledgeway, 1997). According to this hypothesis, any texture contribution to motion at high temporal frequencies would be due to some form of nonlinearity, enabling the first-order motion system to process contrast-defined motion. On the other hand, Lu and Sperling (2001) argued that the absence of texture contribution to motion at high temporal frequencies using a dynamic noise carrier was due to masking, not to a temporally lowpass property of the second-order motion system. To maximize the potential texture contribution to motion at high temporal frequencies, the current study used static noise carriers.

Given that texture can contribute to motion at high temporal frequencies, the question is whether this is due to a second-order motion system or preprocessing nonlinearities introducing global distortion products (Figure 1) processed by the first-order system. If such global distortion products would enable contrast-defined motion processing, then it should be possible to neutralize (or greatly reduce) these distortions by superimposing a luminance modulation of the opposed polarity. However, Badcock and Derrington (1989) were unable to impair the sensitivity to contrast-defined motion by superimposing a luminance modulation, which suggests the existence of another motion system (either feature-tracking or second-order). Scott-Samuel and Georgeson (1999) found different results at high temporal frequencies: Global distortion products can

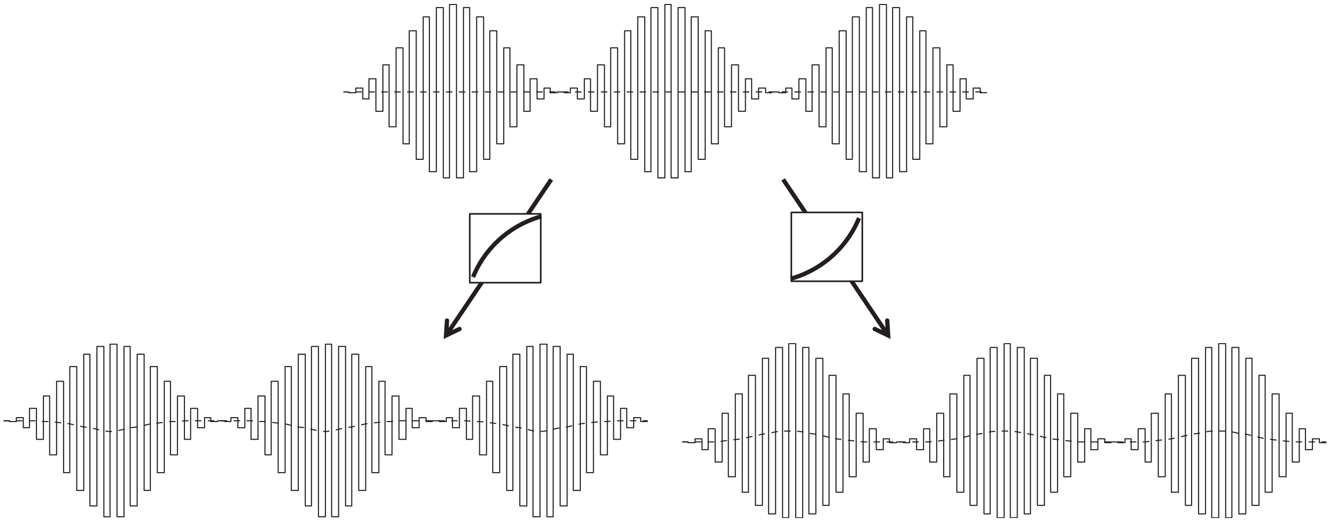


Figure 1. Global distortion product introducing luminance artefacts. The graph on the top shows the luminance profile of a contrast modulation (solid line) in which the mean luminance (dashed line) remains constant. Expansive and compressive nonlinearities would introduce distortion products (i.e., luminance modulations) in phase and in opposite phase, respectively, with the contrast modulation as represented in the bottom right and left, respectively.

be strong enough to enable the first-order motion system to process contrast-defined motion. Nevertheless, nulling these distortion products by introducing an opposite distortion within the stimulus was *not* found to completely neutralize the texture contribution to motion (i.e., contrast-defined motion was still perceptible). They took this as evidence of a second-order motion system operating at high temporal frequencies, but in doing so, they implicitly assumed the preprocessing nonlinearities preceding the luminance motion units were exactly the same for all units at all times. This is a strong assumption. Given that there is internal noise at all processing levels, there must be some variation in the preprocessing nonlinearities of different first-order motion units (and/or across time for the same units). For instance, because some ganglion cells have different equiluminance points (Lee, Martin, & Valberg, 1989), different nonlinearities induced by L and M photoreceptors would result in different preprocessing nonlinearities for different ganglion cells. Consequently, compensating for nonuniform distortion products by introducing a spatiotemporally uniform distortion within the stimulus would leave residual distortion products, namely luminance modulations of different polarities for different motion units (Figure 1). These residual distortion products would be drifting in the same direction as the contrast modulation and could be processed by first-order motion units.

The target of the present study was to determine if the texture contribution to motion at high temporal frequencies is processed by the first-order motion system due to residual distortion products. If it is, it would suggest either that there is no dedicated second-order motion system or that it is temporally lowpass. Conversely, if the texture contribution to motion cannot be explained by residual distortion products, then this would imply the existence of a second-order motion system that can operate at high temporal frequencies. To test this, we adapted a paradigm developed by Cavanagh and Anstis (1991) to test whether the color contribution to motion is due to an equiluminance variability of luminance motion units. The logic of this paradigm is that adding residual distortion products to a high-contrast luminance modulation should have little impact on the contribution to motion. The residual distortion products would sum with the luminance modulation for some units and subtract for others (Figure 2). As a result, residual distortion products would increase the contribution to motion for some first-order motion units and reduce it for others, resulting in no or little additional contribution to motion. The present study therefore evaluated the texture contribution to motion when superimposing a luminance and a contrast modulation having the same spatiotemporal frequency. If the texture contribution to motion is due to residual distortion products, then adding a contrast modulation to a luminance modulation should *not* substantially increase the contribution to motion; that is, the motion contributions of superimposed luminance and contrast modulations should be equal to the motion contribution of the luminance modulation alone. However, if there is a second-order motion system sensitive to high temporal frequencies, we would expect the addition of a

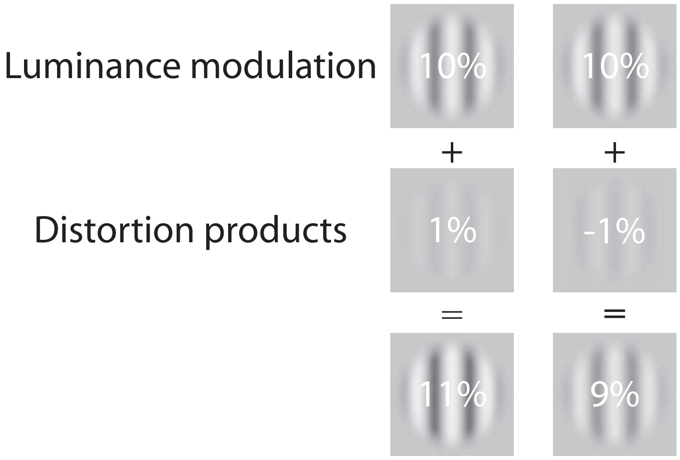


Figure 2. Example of the effect of superimposing residual distortion products and a high-contrast luminance modulation. For some motion units, the distortion products would correspond to a positive luminance modulation (e.g., +1%), and for other motion units, they would correspond to a negative (e.g., −1%), which is in opposite phase. Summing these distortion products with a high-contrast luminance modulation (e.g., 10%) would increase the effective contrast for some motion units (e.g., 11%) and reduce it for others (e.g., 9%), resulting in no net gain on average.

contrast modulation to increase the contribution to motion; that is, the motion contributions of superimposed luminance and contrast modulations should be greater than the motion contribution of the luminance modulation alone.

To measure the contributions to motion of superimposed luminance and contrast modulations, we used a nulling technique, consisting of opposing another luminance modulation drifting in the opposite direction (as done for color by Cavanagh & Anstis, 1991). The contrast of one of the two opposing luminance modulations was manipulated to find the point at which the observer perceived no net motion. The texture contribution to motion was quantified as the contrast difference between the two luminance modulations when no net motion was perceived. This contrast difference corresponds to the extra contribution to motion due to the contrast modulation, i.e., the texture contribution to motion. If the texture does *not* contribute to motion, then the observer should not experience any net motion when the two luminance modulations have the same contrast. Conversely, if there is a motion system other than the first-order motion system sensitive to contrast-defined motion, then the contrast of the opposing luminance modulation would need to be greater than the contrast of the other luminance modulation (the one with the contrast modulation) to cancel both the luminance and contrast modulations.

Experiment 1: Texture contribution to motion at low and high temporal frequencies

The target of the first experiment was to evaluate, at both low and high temporal frequencies, the texture contribution to motion when neutralizing both global and residual distortion products. Because we have at least one motion system other than the first-order motion system responding to contrast-defined motion at low temporal frequencies, the texture contribution to motion at a low temporal frequency should remain substantial even when neutralizing distortion products. At temporal frequencies too high for the feature-tracking motion system, however, we expect a substantial texture contribution to motion only if we have a second-order motion system sensitive to contrast-defined motion at such temporal frequency. If the texture contribution to motion is only due to distortion products, then we expect no substantial texture contribution to motion when neutralizing such artifacts.

Method

Observers

Five naïve observers and one of the authors participated in this experiment. They all had normal or corrected-to-normal vision.

Apparatus

The stimuli were presented on a 19-inch ViewSonic E90FB .25 CRT monitor with a mean luminance of 47 cd/m^2 and a refresh rate of 120 Hz, which was powered by a Pentium 4 computer with a Matrox Parhelia 512 graphics card. The noisy-bit method (Allard & Faubert, 2008b) implemented with the error of the green color gun inversely correlated with the error of the two other color guns made the 8-bit display perceptually equivalent to an analog display having a continuous luminance resolution. The monitor was the only source of light in the room. A Minolta CS100 photometer interfaced with a homemade program calibrated the output intensity of each gun. A control experiment was run to ensure that the display was well calibrated (see results and discussion section below). At the viewing distance of 114 cm, the width and height of each pixel were 1/64° of visual angle.

Stimuli

Stimuli were composed of three sine wave modulations (Figure 3): a luminance and a contrast modula-

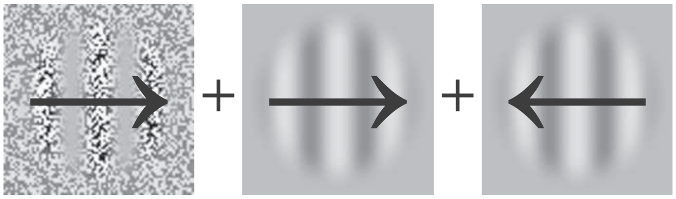


Figure 3. Stimuli consisted of three superimposed modulations: two opposing luminance modulations and a contrast modulation. The texture contribution to motion was defined as the contrast difference between the two luminance modulations when no net motion was perceived, that is, when the contribution to motion of the luminance and contrast modulation drifting in one direction was equal to the luminance modulation drifting in the opposite direction.

tion drifting in the same direction (randomly either left or right) and a luminance modulation drifting in the opposite direction. All three modulations were vertically oriented and had the same spatiotemporal frequency. The spatial frequency was always 0.7 cpd, and the temporal frequency was either 3.75 or 15 Hz. The initial phases of the modulations were randomized, but the luminance and contrast modulations drifting in the same direction were constrained to be superimposed in phase or in opposite phase (i.e., a phase difference of either 0° or 180°). The spatial window was circular with a diameter of 2.8° and soft edges following a half cosine of 0.5°. The presentation time was 250 msec plus an onset and offset half cosine ramp of 125 msec.

Static binary noise with elements of 4 × 4 pixels (i.e., 0.0625 × 0.0625 dva), and 20% contrast was used as a carrier. There was no luminance variation within each noise element to avoid luminance motion drifting cues within noise elements. The carrier was presented over the entire screen and was visible at all times (i.e., it was not modulated by the spatial and temporal windows) to avoid introducing additional distortion products due to the carrier onset (Allard & Faubert, 2008a). A fixation point was continuously presented at the center of the screen.

The contrast of the contrast modulation was always maximized to 100% modulation depth (i.e., low and high contrast strips were of 0% and 40% contrast); the contrast of the luminance modulation drifting with the contrast modulation was the dependent variable, and the contrast of the opposing luminance modulation was systematically varied but was constant within one block.

Procedure

Observers indicated by pressing one of two keys whether they perceived the net direction of the presented stimulus as drifting to the left or to the right. Note that at the low temporal frequency a pilot experiment revealed that observers could perceive motion transparency, so they were asked to judge in which direction the motion was the "strongest." At the high temporal frequency, observers did not report any motion transparency.

A one-down-one-up staircase procedure (Levitt, 1971) was used to control the contrast of the luminance modulation drifting with the contrast modulation in order to converge to the contrast of the luminance modulation at which no net motion was perceived. The contrast of the luminance modulation was reduced or increased depending on whether the net perceived direction was the same or opposite, respectively, as the contrast modulation. Each staircase ended after 200 trials.

For each of the two temporal frequencies (3.75 and 15 Hz), five different contrasts of the opposing luminance modulation were used, resulting in a total of 10 blocks. For each block, two staircases were pseudorandomly interleaved for the two relative phase conditions, that is, luminance and contrast modulations superimposed in phase and in opposite phase.

The main goal of the current experiment was to measure the texture contribution to motion when neutralizing both residual and global distortion products. For comparative reasons, we were also interested in measuring the texture contribution to motion when neutralizing the global but not the residual distortion products. We therefore tested the texture contribution to motion when the contrast of the luminance modulation drifting with the contrast modulation was fixed to compensate for the global distortion product, leaving only the residual distortion products. In this condition, the texture contribution to motion was estimated by manipulating the contrast of the opposing luminance modulation to find the point at which no net motion was perceived. Because there was no luminance modulation drifting with the contrast modulation (besides the one nulling the global distortion product), there was only one staircase within each block, and the estimated texture contribution to motion corresponded to the contrast of the opposing luminance modulation when no net motion was perceived. As detailed in the next section, data analysis consisting of measuring the texture contribution to motion in the presence of a superimposed luminance modulation in phase and opposite phase also provided a measure of the global distortion product for each contrast of the opposing luminance modulation. To estimate the global distortion product when no luminance modulation was superimposed onto the contrast modulation (besides the one nulling the global distortion product), we used the measured global distortion product when the contrast of the opposing luminance modulation was the lowest (i.e., 15% and 5% for the low and high temporal frequencies, respectively). We used this measure of the global distortion product because it was the condition closest to the one without any superimposed luminance modulation.

Data analysis

As presented in the Introduction, nonuniform preprocessing nonlinearities within the visual system can convert contrast modulations into luminance distortion products. Distortion products can be decomposed into *global* distortion products (i.e., a spatiotemporally uniform luminance modulation) and *residual* distortion products (i.e., luminance modulations in phase with the contrast modulation for some motion units and in opposite phase for others). The current paradigm enables the measurement of the texture contribution to motion when neutralizing *both* the global and residual distortion products. As shown in Figure 2, residual distortion products should have little net contribution to motion in the presence of a high-contrast luminance modulation because some would sum with the luminance modulation and the others would subtract. A global distortion product would either sum or subtract with the luminance modulation, depending on their relative phase. Consequently, it would increase the measured contrast of the luminance modulation drifting with the contrast modulation for one relative phase condition (luminance and contrast modulations in phase or in opposite phase) and reduce it, by the same amount, in the other phase condition. As a result, the impact of the global distortion product can be neutralized by averaging the measured contrasts of the luminance modulation at which no net motion is perceived in these two phase conditions. Furthermore, this paradigm also enables the quantification of the global distortion product, which corresponds to half of the difference between the texture contribution to motion when the luminance and contrast modulations are combined in phase and in opposite phase.

For each block of trials (i.e., each temporal frequency and each contrast of the opposing luminance modulation, l_{opp}), the nulling point at which no net motion was perceived was estimated by fitting a Gaussian cumulative distribution function (cdf(μ,σ)) on the log contrast of the luminance modulation drifting with the contrast modulation. The fit had three degrees of freedom: the contrast of the luminance modulation at which no net motion was perceived on average (c), the slope of the cdf (σ), and the global distortion product (g), which was summed or subtracted (in linear units) with the luminance modulation. Consequently, for a given contrast of the opposing luminance modulation (l_{opp}), the probability that the observer perceived the net motion in the direction of the contrast modulation that is summed with a luminance modulation of contrast l either in phase ($\theta = 0$) or in opposite phase ($\theta = 180$) was fitted by maximizing the likelihood of the following function:

$$P_{l_{\text{opp}}}(l,\theta) = \text{cdf}\left(\log\left(\frac{l + p_\theta}{c}\right), \sigma\right) \qquad \begin{matrix} p_0 = -g \\ p_{180} = g \end{matrix}. \tag{1}$$

The texture contribution to motion (t) was quantified as the difference between the contrasts of the two luminance modulations at which no net motion was perceived:

$$t = l_{\text{opp}} - c. \tag{2}$$

To evaluate the texture contribution to motion when there was no luminance modulation ($l = 0$) superimposed to the contrast modulation (except for the luminance modulation nulling the global distortion product, which was fixed for each condition), the data were also fitted with a Gaussian cumulative distribution function:

$$P(l_{\text{opp}}) = 1 - \text{cdf}\left(\log\left(\frac{l_{\text{opp}}}{t}\right), \sigma\right). \tag{3}$$

Results and discussion

At the low temporal frequency (3.75 Hz, Figure 4, left), the texture contribution to motion remained substantial even when a high-contrast luminance modulation was superimposed on the contrast modulation, that is, when both the global and residual distortion products were neutralized. This suggests that the texture contribution to motion at this low temporal frequency cannot be due solely to preprocessing nonlinearities introducing distortion products that are then processed by first-order motion units. We conclude that there is a motion system other than the first-order motion system enabling the texture contribution to motion at this low temporal frequency. This motion system could be a second-order or feature-tracking motion system.

At a temporal frequency too fast for the feature-tracking motion system (15 Hz), we obtained a different pattern of results (Figure 4, right): Superimposing a high-contrast luminance modulation nulled the texture contribution to motion. In other words, under conditions in which the contribution to motion of distortion products should be neutralized and in which the feature tracking cannot operate, the texture did not contribute to motion. These results are consistent with the hypothesis that the texture contribution to motion at high temporal frequencies was due to residual distortion products. This suggests either that

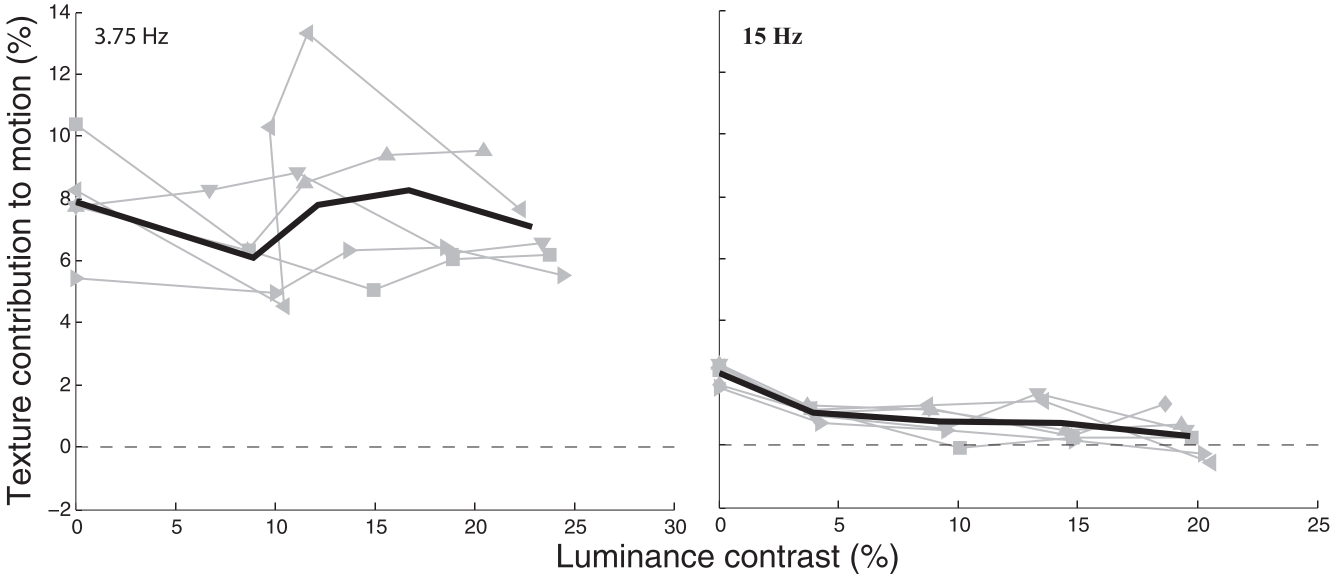


Figure 4. Texture contribution to motion (*t*) at low (3.75 Hz, left) and high (15 Hz, right) temporal frequencies as a function of the contrast of the luminance modulation (*c*). The point at zero luminance contrast corresponds to the condition in which only the global distortion is neutralized. Gray lines and symbols represent the data of individual observers (◀, ▶, ▲, ▼, ■, and ◆ for nbg, njl, md, ra, dlp, and ls, respectively), and the thick black line represents the average across observers. Note that the results of observer ls are not shown at the low temporal frequency (left) because her texture contribution to motion was too high to be measured except for the highest luminance contrast at which the texture contribution to motion was 22%.

there is no dedicated second-order motion system or that it is temporally lowpass.

Figure 5 shows the measured global distortion products. At the high temporal frequency (15 Hz), the global distortion products were generally negative, which is consistent with previous findings reporting early compressive nonlinearities (Scott-Samuel & Georgeson, 1999). Note that to assert that the source of these global distortion products was the visual system and not the display, we ran a control experiment in which an observer's (one of the authors) vision was blurred by wearing positive lenses. The power of the lenses was adjusted so that the visibility at high spatial frequencies was substantially affected, but the luminance motion at low spatial frequencies was still visible. By wearing these lenses, the texture contribution to motion and the global distortion product both dropped near zero at both low and high temporal frequencies. This asserts that the texture contribution to motion and the global distortion product measured were due to the visual system and not to display artifacts.

At the low temporal frequency (3.75 Hz), the global distortion products were near zero for two observers and negative for the other three. The negative global distortion products are surprising given the fact that no substantial nonlinearities are generally observed at low temporal frequencies (Allard & Faubert, 2008a; Scott-Samuel & Georgeson, 1999). Nonetheless, these negative distortion products do not necessarily imply the existence of early compressive nonlinearities at this low temporal frequency. If the motion system enabling the texture contribution to motion was the feature-tracking motion system, it may have been easier to track the combination of the luminance and contrast modulations when they were superimposed in phase rather than in opposite phase for these two observers. Whatever the cause of this measured global distortion product, the important finding that the texture contribution to motion at the low temporal frequency cannot be entirely explained by distortion products remains, which implies that there is a motion system sensitive to low temporal frequencies other than the first-order motion system.

Experiment 2: Avoiding local luminance motion patches

In the previous experiment, we found that at a high temporal frequency superimposing a luminance modulation to a contrast modulation neutralized the texture contribution to motion. This is consistent with the hypothesis that the texture contribution to motion in absence of a luminance modulation was due to residual distortion products (i.e., luminance modulations of different polarities for different motion units) processed by first-order motion units, but it is also consistent with the hypothesis that local texture imbalance introduced local luminance motion patches (Smith & Ledgeway,

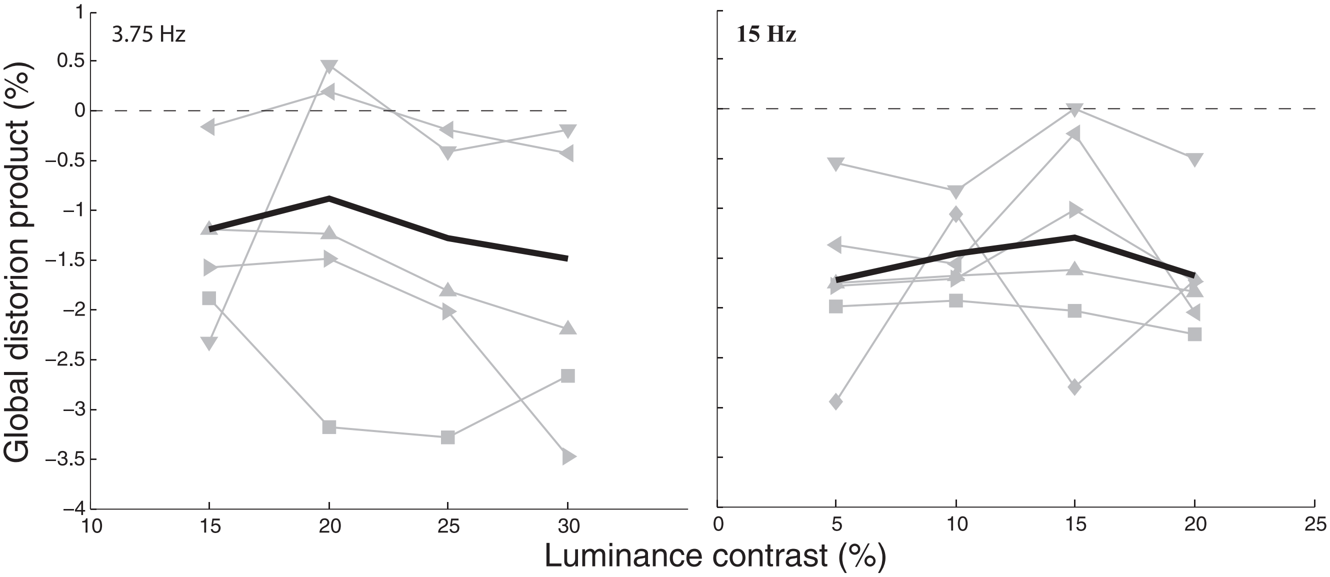


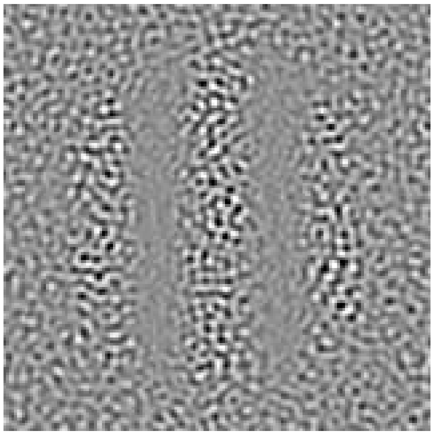
Figure 5. Mean global distortion products (*g*) at low (3.75 Hz, left) and high (15 Hz, right) temporal frequencies as a function of the contrast of the luminance modulation (*c*). Legend as in Figure 4.

1997). For local areas in which the texture is perfectly balanced (e.g., the same proportion of black and white checks), a contrast modulation would not introduce any luminance modulation. Conversely, for unbalanced local areas (e.g., more white or black checks), a contrast modulation would introduce a luminance modulation that is either in phase or in opposite phase, respectively, with the contrast modulation. Thus, the contrast modulation of a locally unbalanced texture would introduce local luminance modulations in phase and in opposite phase with the contrast modulation. The goal of the present experiment was to determine if the texture contribution to motion at high temporal frequencies was caused by nonuniform preprocessing nonlinearities of the visual system or by the fact that the texture was locally unbalanced.

To minimize local texture imbalance, previous studies have often used carriers composed of only high spatial frequencies (as recommended by Smith & Ledgeway, 1997), which have the drawback of introducing motion imbalance in the Fourier domain (i.e., the stimulus is not drift-balanced). At high temporal frequencies, this motion imbalance can result in perceived motion in the *opposite* direction to the drifting of the texture (Scott-Samuel & Georgeson, 1999; Smith & Ledgeway, 1997). Nonetheless, this undesirable side effect can be avoided simply by filtering the texture stimuli *after* the signal is multiplied by the carrier ($f(carrier \times signal)$ (Movie 2), where f is the spatial frequency filter) rather than only filtering the carrier ($f(carrier) \times signal$). In the present experiment, we measured the texture contribution to motion at a high temporal frequency when minimizing, as much as possible, the local texture imbalance.

Method

The method was essentially the same as the one in the previous experiment. Because we were interested in explaining the texture contribution to motion at a temporal frequency that disables the feature-tracking motion system, we only evaluated the texture contribution to motion at 15 Hz. The carrier was composed of static Gaussian noise with elements of 1 × 1 pixels (i.e., 0.016 × 0.016 dva). The contrast of the carrier was fixed so that an unmodulated carrier had an rms Michelson contrast of 20% after the spatial filtering. The filtering was applied after the carrier was multiplied with the signal defining the contrast modulation but before the addition of the luminance



Movie 2. Contrast-defined motion used in the second experiment. Note that the drifting speed was substantially slowed down for the presentation.

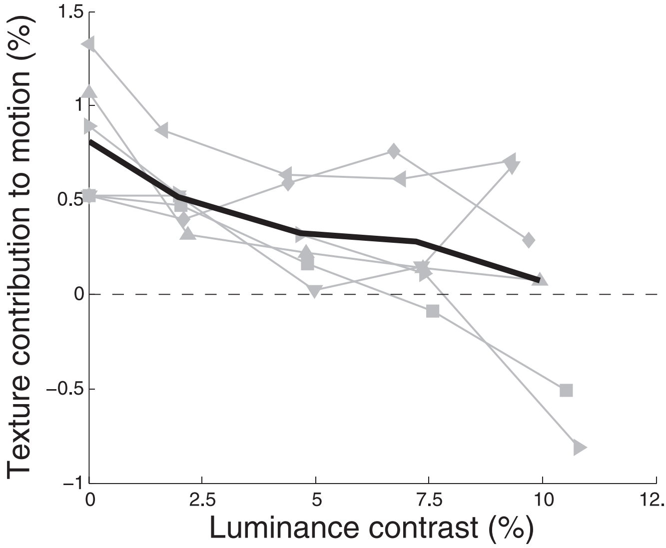


Figure 6. Texture contribution to motion at a high (15 Hz) temporal frequency as a function of the contrast of the luminance modulation when neutralizing both the global and residual distortion products. Legend as in Figure 4.

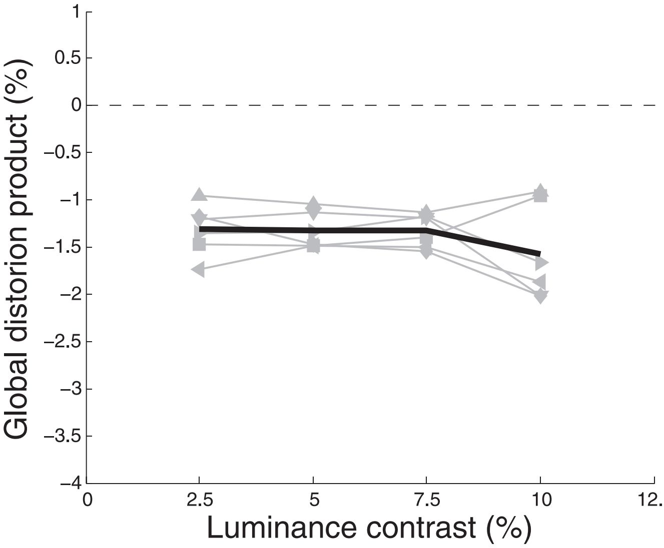


Figure 7. Mean global distortion products (*g*) at a high (15 Hz) temporal frequency as a function of the contrast of the luminance modulation (*c*) when minimizing local texture imbalance. Legend as in Figure 4.

modulations (otherwise they would be filtered out). The spatial filtering was done in the Fourier domain to keep only the spatial frequencies between 4 and 8 cpd. All the other settings were the same as in the previous experiment.

Results and discussion

The pattern of results (Figure 6) was similar to the one observed in the first experiment at the high temporal frequency; the texture contributed to motion when the contrast of the luminance modulation was low but not when it was high. Again, this suggests that there was no second-order motion system operating at this high temporal frequency (15 Hz). Note that there is one observer (njl) who seemed to consistently have a texture contribution greater than zero. To test if this texture contribution to motion was significantly higher than zero, we retested this observer at the highest luminance contrast, and her texture contribution to motion happened to be slightly below zero (−0.11). Furthermore, this observer had a texture contribution to motion that also fell below zero in the previous experiment. We concluded that the apparently consistent texture contribution to motion only for this observer and in the current experiment was due to variance. The global trend of the six observers is clear: The texture contribution to motion dropped near zero when increasing the luminance contrast.

In the absence of a luminance modulation, the texture contribution to motion when the stimulus was filtered to keep only high spatial frequencies was about half of the one using a broadband carrier (Experiment 1). If the texture contribution to motion in Experiment 2 was only due to local texture imbalance, this difference should have been much more. To illustrate this, we conducted simulations to measure the relative residual distortion products due to local texture imbalance when using both types of contrast-modulated stimuli (Experiments 1 and 2). Specifically, we measured the correlation between a vertically oriented luminance Gabor patch having the same modulation as the contrast modulation (0.7 cpd) with various frame samples of contrast-modulated stimuli in both experiments. Given that there was no luminance modulation, the correlation was, on average, near zero in both cases. But the rms of the correlations across many samples (which would correspond to residual distortion products) was more than 100 times greater with the stimuli used in Experiment 1. Note that this was true for Gabors having various sizes (standard deviations of 0.25° and higher). This confirms that the texture contribution to motion was not only due to local texture imbalance because the data showed a difference of a factor of only two. We concluded that the texture contribution to motion was due to residual distortion products caused by preprocessing nonlinearities occurring within the visual system, not to local texture imbalances.

Figure 7 shows the global distortion product, which was, on average, similar to the one observed in the previous experiment.

Experiment 3: Phase dependence test

In the first two experiments, we found that at a high temporal frequency texture did *not* contribute to motion when distortion products were neutralized by the superimposition of a high-contrast luminance modulation. It has been argued (Lu & Sperling, 1995, 2001) that if contrast modulations were converted into luminance modulations (i.e., distortion products) due to preprocessing nonlinearities, then the texture contribution to motion should depend on the relative phase between the luminance and contrast modulations. This is certainly true for a texture contribution to motion due to *global* distortion products, which would either sum or subtract with a luminance modulation, depending on whether the luminance and contrast modulations were combined in phase or in opposite phase (i.e., differed by 0° and 180°). However, a texture contribution to motion due to *residual* distortion products would not necessarily depend on the phase difference between the luminance and contrast modulations. For instance, according to the energy model (Adelson & Bergen, 1985), the motion responses of first-order motion units would be proportional to the motion energy (i.e., the square of the luminance contrast), so we would expect no phase interaction (the energy sum of two uncorrelated signals is equal to the energy of the combined signals). Because residual distortion products (which correspond to luminance modulations of opposite polarities) are not correlated with a uniform luminance modulation, then the energy sum of these two combined modulations would always be the same, independent of their phase difference (i.e., the energy of residual distortion products plus the energy of the luminance modulation). Thus, if the texture contribution to motion is due to residual distortion products and if the motion responses of first-order motion units are proportional to the motion energy (i.e., squared contrast), then the texture contribution to motion should be independent of the phase difference between the luminance and contrast modulations.

Nonetheless, although the absence of phase interaction between luminance- and contrast-defined motion processing would not necessarily imply distinct processing, a phase interaction would imply some common processing. A texture contribution to motion due to residual distortion products would predict a phase interaction if the motion responses of first-order motion units were *not* proportional to the squared luminance contrast (i.e., energy). We therefore sought such a phase interaction. For instance, if the motion unit responses were proportional to the luminance contrast, then the texture contribution to motion would be maximized when the phase of the luminance and contrast modulations differed by 90° (i.e., a 90° or 270° difference). But if the motion unit response were more expansive than the quadratic function (e.g., thresholding mechanisms), then the texture contribution to motion would be maximized when the phase of the luminance and contrast modulations differed by 0° or 180°. In the current experiment, we therefore evaluated the texture contribution to motion when the phase of the superimposed luminance and contrast modulations differed by 0°, 90°, 180°, and 270°.

Method

The method was highly similar to the one used in the previous experiment. It was important to evaluate the phase interaction under conditions in which the texture contribution to motion was *not* due to local texture imbalances. Indeed, a phase interaction when the texture contribution to motion was not due to local texture imbalances would imply the existence of preprocessing nonlinearities enabling at least some texture contributions to motion. We therefore filtered the texture-defined stimuli to keep only high spatial frequencies (i.e., between 4 and 8 cpd) as done in the previous experiment. Obviously, a phase interaction is expected to be weak when either the luminance or contrast modulation does *not* substantially contribute to motion. To equate the texture and luminance contributions to motion, the contrast of the luminance modulation was fixed to the texture contribution to motion in the absence of a luminance modulation ($l = 0$) observed in Experiment 2 (Figure 6, i.e., 0.0058, 0.0133, 0.0089, 0.0107, 0.0052, and 0.0052 for observers nbg, njl, md, ra, dlp, and ls, respectively). To neutralize the global distortion product, another luminance modulation drifting in phase with the contrast modulation was also superimposed. The contrast of this luminance modulation was estimated by a pilot experiment in order to equate the texture contribution to motion when the luminance and contrast modulations were combined in phase and in opposite phase (0.0101, 0.0109, 0.0098, 0.0122, 0.0146, and 0.0136 for observers nbg, njl, md, ra, dlp, and ls, respectively). To measure the texture contribution to motion of these combined modulations, we opposed another luminance modulation drifting in the opposite direction. The contrast of this opposing luminance modulation was manipulated for each phase difference by a staircase procedure. The four staircases (i.e., 0°, 90°, 180°, and 270° difference) ended after 200 trials and were interlaced within one block.

Results and discussion

Figure 8 presents the texture contribution to motion as a function of the phase difference between the

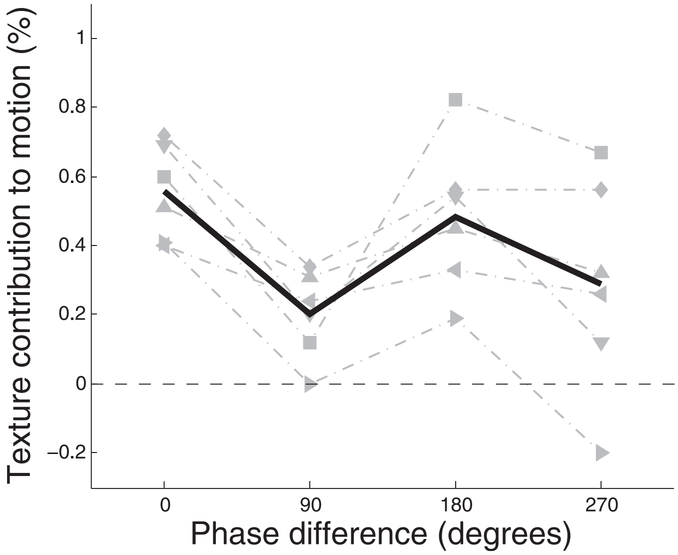


Figure 8. Texture contribution to motion at a high (15 Hz) temporal frequency as a function of the phase difference between the luminance and contrast modulations when neutralizing the global distortion product and minimizing local carrier imbalance. Legend as in Figure 4.

luminance and contrast modulations. As expected, the texture contributions to motion were similar whether the luminance and contrast modulations were superimposed in phase (0°) or in opposite phase (180°). This confirms that the global distortion products were neutralized. As would be expected, the texture contributions to motion were also similar when the luminance and contrast modulations differed by ±90° (i.e., 90° and 270°). More importantly, the texture contribution to motion depended on the phase interaction between the luminance and contrast modulations: The texture contributions to motion were more important when the phase difference between the luminance and contrast modulations were 0° or 180° than 90° or 270° ($p < .01$). This phase interaction implies at least some common processing. Because the phase information is lost when extracting the motion, this interaction must occur *before* the motion-extraction stage. This phase interaction is compatible with the hypothesis that the texture contribution to motion was due to residual distortion products and, more importantly, cannot be explained by distinct luminance- and contrast-defined motion processing.

The fact that the texture contribution to motion was greater when the phases of the luminance and contrast modulations differed by 0° or 180° compared to 90° or 270° suggests that the relation between the luminance contrast and the motion response of first-order motion units was more expansive than a quadratic function.

General discussion

The first experiment showed that superimposing a luminance modulation onto the contrast modulation neutralized the texture contribution to motion at high temporal frequencies but not at low temporal frequencies. The second experiment showed that the texture contribution to motion at high temporal frequencies (when it was not completely neutralized by the luminance modulation) was not due to luminance motion patches caused by local texture imbalances. In the third experiment, we found that the texture contribution to motion depended on the phase difference between the luminance and contrast modulations, which implies at least some common processing before the motion-extraction stage.

The substantial texture contribution to motion at a low temporal frequency when the distortion products were neutralized confirms that there must be a motion system other than the luminance system sensitive to contrast-defined motion. These results are consistent with the substantial amount of evidence showing separate pathways for luminance- and contrast-defined motion processing at low temporal frequencies, which implies the existence of a motion system sensitive to contrast-defined motion that is relatively unaffected by the superimposition of a luminance modulation. This is consistent with the fact that superimposing luminance-modulated noise interfered with contrast-defined motion processing at high but not at low temporal frequencies (Allard & Faubert, 2008a). In the current study, because the texture contribution to motion dropped near zero at high temporal frequencies when superimposing a luminance modulation, the motion system sensitive to contrast-defined motion and relatively unaffected by the superimposition of a luminance modulation did not operate at high temporal frequencies. This motion system could either be the feature-tracking motion system or a dedicated second-order motion system that is temporally lowpass.

Note that other studies have shown that the motion system that is most sensitive to contrast-defined motion can differ at low and high temporal frequencies (Holliday & Anderson, 1994; Seiffert & Cavanagh, 1999; Ukkonen & Derrington, 2000). However, the results of these studies do not imply that the most sensitive motion system operating at low temporal frequencies does not contribute to motion at high temporal frequencies as it could still contribute to motion without being the *most sensitive*. Here, the complete neutralization of the texture contribution to motion at high temporal frequencies by superimposing a luminance modulation shows that there is no motion system operating at high temporal frequencies that is unaffected by the superimposition of a luminance modulation.

The neutralization of the texture contribution to motion at high temporal frequencies (Experiments 1 and 2) shows that residual distortion products *can* explain the texture contribution to motion. However, this, by itself, does not *necessarily* imply that the texture contribution to motion was not due to a second-order motion system. Indeed, a nonlinear combination of the outputs of first- and second-order motion units (e.g., a sum of squares) could also explain that the texture contribution to motion drops when superimposing a luminance modulation. Nonetheless, the phase interactions between the processing of superimposed luminance and contrast modulations observed in Experiment 3 imply some common processing *before* the motion-extraction stage, which rules out the possibility that the texture contribution to motion at high temporal frequencies was *only* due to a second-order motion system. As a result, there must have been at least a large proportion (if not all) of the texture contribution to motion at high temporal frequencies that was due to residual distortion products.

Furthermore, this phase interaction cannot be explained by the gradient-based model proposed by Benton, Johnston, and colleagues (Benton, 2002, 2004; Benton & Johnston, 2001; Benton, Johnston, McOwan, & Victor, 2001; Johnston & Clifford, 1995; Johnston, McOwan, & Buxton, 1992). The keystone of their model is the calculation of the temporal derivative relative to the spatial derivative. Based on computational simulations, they showed that applying a nonlinearity after such a calculation could extract the drifting direction of contrast-defined motion. Because this model proposes the computation of the spatial and temporal derivative of the stimulus, the mechanism processing such a stimulus should be tuned to spatial frequencies corresponding to the spectral component of the stimulus, which, for contrast-defined motion, directly depends on the spatial frequency of the carrier and not on the spatial frequency of the envelope, which requires a preprocessing nonlinearity to be recovered. Consequently, the gradient-based model would predict no phase interaction between luminance- and contrast-defined motion processing because the motion energy of both types of motion would be distributed across different frequency ranges (0.7 cpd of the luminance modulation and between 4 and 8 cpd for the contrast modulation). Consequently, the phase interaction observed in Experiment 3 suggests that the texture contribution to motion was due to preprocessing nonlinearities that introduced luminance-defined motion of the same spatiotemporal frequency as the contrast modulation.

Our conclusion that the texture contribution to motion at high temporal frequencies was due to the first-order motion system supports the argument of Ledgeway and colleagues (Hutchinson & Ledgeway, 2006; Smith & Ledgeway, 1997) that the texture contribution to motion observed at high temporal frequencies with static noise carriers was due to the first-order motion system and not to a dedicated second-order motion system as concluded by others (Lu & Sperling, 1995, 2001; Scott-Samuel & Georgeson, 1999). However, Ledgeway and colleagues argued that the first-order motion system was able to process contrast-defined motion due to local texture imbalances; whereas we showed that the texture contribution to motion was due to nonuniform preprocessing nonlinearities within the visual system. Even when filtering the stimulus to keep only high spatial frequencies (which should remove nearly all local texture imbalances), the texture contribution to motion remained substantial (Experiment 2). Thus, the current study suggests the existence of nonuniform preprocessing nonlinearities that can enable some texture contribution to motion and that have not been considered in previous studies. It will be interesting to investigate whether such nonlinearities could explain some texture contribution to motion at low temporal frequencies that have been attributed to a dedicated second-order motion system.

Note that further evidence of these residual distortion products comes from the fact that they can be seen. Harry Orbach mentioned to us (personal communication at VSS 2008) that when viewing a flickering Ganzfeld at a high contrast and a high temporal frequency, he observed that the field appeared spatially nonuniform even though it was. After this conversation, we created a large flickering stimulus, and we observed the same phenomenon. More specifically, the greater the contrast and the higher the temporal frequency, the less uniform the flickering field appeared. This subjective experience directly demonstrates the existence of residual distortion products at high temporal frequencies.

In the current study, we found a substantial texture contribution to motion that cannot be due to the first-order motion system only at a low temporal frequency at which the feature-tracking motion system can operate (3.75 Hz). Thus, our results are consistent with studies suggesting that attention is involved in contrast-defined motion processing (Ashida et al., 2001; Derrington & Ukkonen, 1999; Ukkonen & Derrington, 2000). On the other hand, attributing contrast-defined motion processing only to the first-order and feature-tracking motion systems seems incompatible with other studies suggesting that luminance- and contrast-defined motions are processed by distinct motion systems even for stimuli that presumably cannot be processed by the feature-tracking motion system, such as random-dot patterns (Badcock & Khuu, 2001; Cassanello, Edwards, Badcock, & Nishida, 2011; Edwards & Badcock, 1995; Mather & West, 1993). However, the idea that luminance- and contrast-defined motions are processed by distinct motion systems even for high-level motion integration (global motion) is controversial (Aaen-Stockdale, Ledgeway, McGraw, & Hess, 2012). Further investigations will be required to clarify the role of attention in contrast-defined motion processing (when it is not due to the first-order motion system).

Given that the texture contribution to motion was weak at high temporal frequencies, that it could be neutralized by the superimposition of a luminance modulation, and that the visual system introduces at least some residual distortion products resulting in at least a large proportion of the texture contribution to motion, we find no reason to suggest the existence of a second-order motion system enabling some texture contribution to motion at high temporal frequencies. We conclude that the texture contribution to motion at high temporal frequencies was due to nonuniform preprocessing nonlinearities within the visual system enabling first-order motion units to process distortion products and not to a second-order motion system.

Keywords: motion, second-order, distortion product, high temporal frequency, contrast-defined motion

Acknowledgments

This research was supported by NSERC discovery fund and by the NSERC-Essilor industrial research chair awarded to Jocelyn Faubert.

Commercial relationships: none.

Corresponding author: Rémy Allard.
Email: remy.allard@umontreal.ca.
Address: Visual Psychophysics and Perception Laboratory, Université de Montréal, Canada.

References

Aaen-Stockdale, C., Ledgeway, T., McGraw, P., & Hess, R. F. (2012). Interaction of first- and second-order signals in the extraction of global-motion and optic-flow. *Vision Research, 68,* 28–39, doi:10.1016/j.visres.2012.07.004.

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

Allard, R., & Faubert, J. (2008a). First- and second-order motion mechanisms are distinct at low but common at high temporal frequencies. *Journal of Vision, 8*(2):12, 1–17, http://www.journalofvision.org/content/8/2/12, doi:10.1167/8.2.12. [PubMed] [Article].

Allard, R., & Faubert, J. (2008b). The noisy-bit method for digital displays: Converting a 256 luminance resolution into a continuous resolution. *Behavior Research Methods*, *40*(3), 735–743.

Anstis, S. M., & Mackay, D. M. (1980). The perception of apparent movement [and discussion]. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences, 290*(1038), 153–168, doi:10.1098/rstb.1980.0088.

Ashida, H., Seiffert, A. E., & Osaka, N. (2001). Inefficient visual search for second-order motion. *Journal of the Optical Society of America, A, Optics, Image Science, & Vision*, *18*(9), 2255–2266.

Badcock, D. R., & Derrington, A. M. (1989). Detecting the displacements of spatial beats: No role for distortion products. *Vision Research, 29*(6), 731–739.

Badcock, D. R., & Khuu, S. K. (2001). Independent first- and second-order motion energy analyses of optic flow. *Psychological Research, 65*(1), 50–56.

Benton, C. P. (2002). Gradient-based analysis of non-Fourier motion. *Vision Research*, *42*(26), 2869–2877.

Benton, C. P. (2004). A role for contrast-normalisation in second-order motion perception. *Vision Research*, *44*(1), 91–98.

Benton, C. P., & Johnston, A. (2001). A new approach to analysing texture-defined motion. *Proceedings of the Royal Society of London Series B: Biological Sciences*, *268*(1484), 2435–2443.

Benton, C. P., Johnston, A., McOwan, P. W., & Victor, J. D. (2001). Computational modeling of non-Fourier motion: Further evidence for a single luminance-based mechanism. *Journal of the Optical Society of America, A, Optics, Image Science, & Vision, 18*(9), 2204–2208.

Braddick, O. J. (1974). A short-range process in apparent motion. *Vision Research, 14*(7), 519–527, doi:10.1016/0042-6989(74)90041-8.

Braddick, O. J., Ruddock, K. H., Morgan, M. J., & Marr, D. (1980). Low-level and high-level processes in apparent motion [and discussion]. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 290*(1038), 137–151.

Burr, D., & Thompson, P. (2011). Motion psychophysics: 1985–2010. *Vision Research, 51*(13), 1431–1456, doi:10.1016/j.visres.2011.02.008.

Cassanello, C. R., Edwards, M., Badcock, D. R., & Nishida, S. (2011). No interaction of first- and second-order signals in the extraction of global-motion and optic-flow. *Vision Research, 51*(3), 352–361, doi:10.1016/j.visres.2010.11.012.

Cavanagh, P. (1992). Attention-based motion perception. *Science*, *257*(5076), 1563–1565.

Cavanagh, P., & Anstis, S. (1991). The contribution of color to motion in normal and color-deficient observers. *Vision Research, 31*(12), 2109–2148.

Derrington, A. M., Badcock, D. R., & Henning, G. B. (1993). Discriminating the direction of second-order motion at short stimulus durations. *Vision Research, 33*(13), 1785–1794.

Derrington, A. M., & Ukkonen, O. I. (1999). Second-order motion discrimination by feature-tracking. *Vision Research*, *39*(8), 1465–1475.

Edwards, M., & Badcock, D. R. (1995). Global motion perception: No interaction between the first- and second-order motion pathways. *Vision Research, 35*(18), 2589–2602.

Harris, L. R., & Smith, A. T. (1992). Motion defined exclusively by second-order characteristics does not evoke optokinetic nystagmus. *Visual Neuroscience*, *9*(6), 565–570.

Holliday, I. E., & Anderson, S. J. (1994). Different processes underlie the detection of second order motion at low and high temporal frequencies. *Proceedings of the Royal Society of London Series B: Biological Sciences, 257*(1349), 165–173.

Hutchinson, C. V., & Ledgeway, T. (2006). Sensitivity to spatial and temporal modulations of first-order and second-order motion. *Vision Research*, *46*(3), 324–335.

Johnston, A., & Clifford, C. W. (1995). Perceived motion of contrast-modulated gratings: Predictions of the multi-channel gradient model and the role of full-wave rectification. *Vision Research, 35*(12), 1771–1783.

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 Series B: Biological Sciences*, *250*(1329), 297–306.

Julesz, B. (1971). *Foundations of cyclopean perception*. Chicago: University of Chicago Press.

Ledgeway, T., & Hutchinson, C. V. (2005). The influence of spatial and temporal noise on the detection of first-order and second-order orientation and motion direction. *Vision Research*, *45*(16), 2081–2094.

Lee, B. B., Martin, P. R., & Valberg, A. (1989). Nonlinear summation of M- and L-cone inputs to phasic retinal ganglion cells of the macaque. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience, 9*(4), 1433–1442.

Levitt, H. (1971). Transformed up-down methods in psychoacoustics. *Journal of the Acoustical Society of America, 49*(2), Suppl 2:467–477.

Lu, Z.-L., & Sperling, G. (1995). The functional architecture of human visual motion perception. *Vision Research*, *35*(19), 2697–2722.

Lu, Z.-L., & Sperling, G. (2001). Three-systems theory of human visual motion perception: Review and update. *Journal of the Optical Society of America, A, Optics, Image Science, & Vision, 18*(9), 2331–2370.

Mather, G., & West, S. (1993). Evidence for second-order motion detectors. *Vision Research, 33*(8), 1109–1112.

Nishida, S. (2011). Advancement of motion psychophysics: Review 2001–2010. *Journal of Vision, 11*(5):11, 1–53, http://www.journalofvision.org/content/11/5/11, doi:10.1167/11.5.11. [PubMed] [Article]

Nishida, S., & Sato, T. (1995). Motion aftereffect with flickering test patterns reveals higher stages of motion processing. *Vision Research*, *35*(4), 477–490.

Scott-Samuel, N. E., & Georgeson, M. A. (1999). Does early non-linearity account for second-order motion? *Vision Research*, *39*(17), 2853–2865.

Seiffert, A. E., & Cavanagh, P. (1999). Position-based motion perception for color and texture stimuli: Effects of contrast and speed. *Vision Research, 39*(25), 4172–4185.

Smith, A. T., & Ledgeway, T. (1997). Separate detection of moving luminance and contrast modulations: Fact or artifact? *Vision Research, 37*(1), 45–62.

Ukkonen, O. I., & Derrington, A. M. (2000). Motion of contrast-modulated gratings is analysed by different mechanisms at low and at high contrasts. *Vision Research*, *40*(24), 3359–3371.

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

van Santen, J. P., & Sperling, G. (1985). Elaborated Reichardt detectors. *Journal of the Optical Society of America. A, Optics and Image Science, 2*(2), 300–321.

Watson, A. B., & Ahumada, A. J., Jr. (1985). Model of human visual-motion sensing. *Journal of the Optical Society of America. A, Optics and Image Science*, *2*(2), 322–341.