

Non-Fourier Motion in The Fourier Spectrum

Steven S. Beauchemin

GRASP Laboratory

Department of Computer and Information Science

University of Pennsylvania

Philadelphia PA 19104-6228

Abstract *Typically, image motion analysis in the frequency domain is performed according to the Motion From Fourier Coefficients, or MFFC principle. However, this principle excludes a class of motion that does not generate power spectra containing the origin of the frequency domain, which is usually referred to as Non-Fourier motion. This type of motion includes phenomena relevant to motion analysis such as translucency, sinusoidal beats, occlusion, Theta motion, and lacks a definitive theoretical framework. We address this problem by deriving exact mathematical expressions for Non-Fourier motion, thus providing a basis for computational models for velocity extraction in the frequency domain. Numerical experiments demonstrating the validity of the approach are also presented.*

Keywords: Image motion, Optical flow, Fourier transform, Non-Fourier motion

1 Introduction

Traditionally, motion perception has been equated with orientation of power in the frequency domain. The many optical flow methods use what Chubb and Sperling term the Motion-From-Fourier-Components (MFFC) principle [1] in which the orientation of the plane or line through the origin of the frequency space that contains most of the spectral power gives the rate of image translation.

The MFFC principle states that for a moving stimulus, its Fourier transform has substantial power over some regions of the frequency domain whose points spatiotemporally correspond to sinusoidal gratings with drift direction consonant with the perceived motion [1]. In addition, current models of human perception involve some frequency analysis of the imagery, such as band-pass filtering and similar processes. However, some classes of moving stimuli which elicit a strong percept in subjects fail to show a coherent spatiotemporal frequency distribution of their power and cannot be un-

derstood in terms of the MFFC principle.

Examples include drift-balanced visual stimuli [1], Fourier and Non-Fourier plaid superpositions [4], amplitude envelopes, sinusoidal beats and various multiplicative phenomena [2]. By drift-balanced it is meant that a visual stimulus with two (leftward and rightward, for example) or more different motions shows identical contents of Fourier power for each motion and therefore, according to the MFFC principle, should not elicit a coherent motion percept. However, some classes of drift-balanced stimuli defined by Chubb and Sperling do elicit strong coherent motion percepts, contrary to the predictions of the usual MFFC model.

Sources of Non-Fourier motion also include the motion of texture boundaries and the motion of motion boundaries. For instance, transparency as considered by Fleet and Langley [2] is an example of Non-Fourier motion, as transparency causes the relative scattering of Fourier components away from the spectrum of the moving stimuli. In addition, occlusion is another example of Non-Fourier motion which is closely related to the Theta motion stimuli of Zanker [5], where the occlusion window moves independently from both the foreground and the background, thus involving three independent velocities.

It has been observed by Fleet and Langley that many Non-Fourier motion stimuli have simple characterizations in the frequency domain, namely power distributions located along lines or planes which do not contain the origin of the frequency space, as required by the MFFC idealization [2]. We develop exact frequency representations for several non-Fourier motions and state their properties with respect to image motion.

1.1 Methodology

To analyze the frequency structure of image signals while preserving representations that are as general as possible, an effort is made to only pose those hypotheses that preserve the generality of the analysis to follow. We describe the assumptions and the techniques with

which the theoretical results were obtained.

Image Signals The geometry of visual scenes under perspective projection generally yields complex image signals. Conceptually, assumptions concerning scene structure should not be made, as they constrain the geometry of observable scenes. In addition, any measured physical signal, such as image intensities, satisfies Dirichlet conditions. Such signals admit a finite number of finite discontinuities, are absolutely integrable and may be expanded into complex exponential series. Dirichlet conditions constitute the sum of assumptions made on image signals.

Mathematical Technique The results established in this analysis emanate from a general approach to modeling visual scenes exhibiting non-Fourier motion. In the case of occlusion, an equation which describes the spatio-temporal pattern of the superposition of a background and an occluding signal is established, in which a characteristic function describing the position of an occluding signal within the imaging space of the visual sensor is defined:

$$\chi(\mathbf{x}) = \begin{cases} 1 & \text{if } \mathbf{x} \text{ within occluding signal} \\ 0 & \text{otherwise.} \end{cases} \quad (1)$$

Two image signals $\mathbf{I}_1(\mathbf{x})$ and $\mathbf{I}_2(\mathbf{x})$, corresponding to the occluding and occluded signals respectively, are defined to form the complete signal pattern:

$$\begin{aligned} \mathbf{I}(\mathbf{x}, t) &= \chi(\mathbf{v}_1(\mathbf{x}, t))\mathbf{I}_1(\mathbf{v}_1(\mathbf{x}, t)) \\ &+ [1 - \chi(\mathbf{v}_1(\mathbf{x}, t))]\mathbf{I}_2(\mathbf{v}_2(\mathbf{x}, t)), \end{aligned} \quad (2)$$

where $\mathbf{v}_i(\mathbf{x}, t)$ is velocity. Alternatively, scenes of translucency are modeled as

$$\mathbf{I}(\mathbf{x}, t) = f(\rho_1)(\mathbf{v}_1(\mathbf{x}, t))\mathbf{I}_2(\mathbf{v}_2(\mathbf{x}, t)), \quad (3)$$

where $f(\rho_1)$ is a function of the density of the translucent material.

Hypotheses made on the components of (2) and (3) are inserted and the structure of the corresponding non-Fourier motions in frequency space are developed. That is to say, signal structures are expanded into complex exponential series, such as:

$$\mathbf{I}_i(\mathbf{x}) = \sum_{\mathbf{n}=-\infty}^{\infty} c_{i\mathbf{n}} e^{i\mathbf{x}^T N\mathbf{k}_i}, \quad (4)$$

where $\mathbf{I}_i(\mathbf{x})$ is the i^{th} intensity pattern, $c_{i\mathbf{n}}$ are complex coefficients, \mathbf{k}_i are fundamental frequencies, $\mathbf{n}^T = (n_1, n_2, \dots, n_n)$ are integers and $N = \mathbf{n}^T I$.

Relevance of Fourier Analysis It has been conjectured that Non-Fourier spectra have mathematically simple characterizations in Fourier space [2]. Consequently, the use of Fourier analysis as a local tool is justified as long as one realizes that it constitutes a global idealization of local phenomena. In that sense, Fourier analysis is used as a local tool whenever Gabor filters, wavelets or local Discrete Fourier Transforms are employed for signal analysis.

Experimental Technique Given the theoretical nature of this research, the purpose of the numerical experiments is to verify the validity of the theoretical results. In order to accomplish this, the frequency content of the image signals used in the experiments must be entirely known to the experimenter, thus forbidding the use of natural image sequences. In addition, image signals with single frequency components are used in order to facilitate the interpretation of experiments involving 3D Fast Fourier transforms. The use of more complex signals impedes a careful examination of the numerical results and do not extend the understanding of the phenomena under study in any particular way.

2 Sinusoidal Beats and Translucency

Sinusoidal beats are sums of sinusoidal patterns, each moving with a possibly different velocity, and constitute one of the simplest forms of Non-Fourier motion. A simple beat involving two 1D sinusoids can be expressed as

$$\mathbf{I}(x, t) = \cos(k_1 x + \omega_1 t) + \cos(k_2 x + \omega_2 t), \quad (5)$$

which is recognized as a case of additive translucency involving two patterns. The Fourier transform of sinusoidal beats is straightforward yet does not comply with the MFFC principle as the frequencies of each pattern do not align to represent an unambiguous motion pattern. These types of stimuli have been used to study velocity detection thresholds and lead to the hypothesis that the human visual system may use a non-Fourier channel of motion detection [3].

Figure 1 shows a superposition of two 1D sinusoids and the corresponding Fourier transform. In the context of sinusoidal beats, Δk and $\Delta \omega$ are beat frequencies and define group velocity as $\frac{-\Delta \omega}{\Delta k}$. Average frequencies \bar{k} and $\bar{\omega}$ are carrier frequencies and define carrier velocity as $\frac{\bar{\omega}}{\bar{k}}$.

In the case of additive translucency, we express the translucent event as (3) and assume a spatially constant

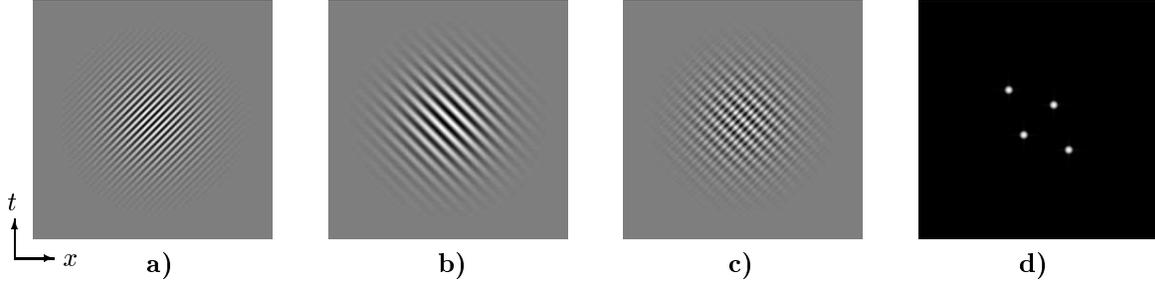


Figure 1: *The composition of an additive transparency scene. a): First sinusoidal signal with frequency $k_1 = \frac{2\pi}{8}$ and velocity $a_1 = 1.0$. b): Second sinusoidal signal with frequency $k_2 = \frac{2\pi}{16}$ and velocity $a_2 = -1.0$. c): Transparency created with the superposition of first and second sinusoidal signal. d): Frequency spectrum of transparency.*

$f(\rho_1)$ with translucency factor φ , leading to a weighted superposition of intensity patterns, written as

$$\mathbf{I}(\mathbf{x}, t) = \varphi \mathbf{I}_1(\mathbf{v}_1(\mathbf{x}, t)) + (1 - \varphi) \mathbf{I}_2(\mathbf{v}_2(\mathbf{x}, t)), \quad (6)$$

where $\mathbf{I}_1(\mathbf{v}_1(\mathbf{x}, t))$ is the intensity profile of the translucent material and $\mathbf{I}_2(\mathbf{v}_2(\mathbf{x}, t))$ is the intensity profile of the background. With $\mathbf{I}_1(\mathbf{v}_1(\mathbf{x}, t))$ and $\mathbf{I}_2(\mathbf{v}_2(\mathbf{x}, t))$ satisfying Dirichlet conditions, the frequency spectrum of (6) is written as

$$\begin{aligned} \hat{\mathbf{I}}(\mathbf{k}, \omega) = & \\ & \varphi \sum_{\mathbf{n}=-\infty}^{\infty} c_{1\mathbf{n}} \delta(\mathbf{k} - N\mathbf{k}_1, \omega + \mathbf{a}_1^T N\mathbf{k}_1) + \\ & (1 - \varphi) \sum_{\mathbf{n}=-\infty}^{\infty} c_{2\mathbf{n}} \delta(\mathbf{k} - N\mathbf{k}_2, \omega + \mathbf{a}_2^T N\mathbf{k}_2). \end{aligned} \quad (7)$$

In this case, the velocities of interest are those of the individual signals. Each of them is predicted by the MFCC principle but the sum of their frequencies does not allow a direct measurement.

3 Occlusion

At an occlusion, the occluding and occluded velocities can always be identified as such and, a degenerate occluding signal exhibiting a linear spectrum is supplemented by the linear orientation of its occluding boundary, allowing to determine the full velocity of the occluding signal.

Let $\mathbf{I}_1(\mathbf{x})$ and $\mathbf{I}_2(\mathbf{x})$ be 2D functions satisfying Dirichlet conditions such that they may be expressed as complex exponential series expansions (3). Also let $\mathbf{I}_1(\mathbf{x}, t) = \mathbf{I}_1(\mathbf{v}_1(\mathbf{x}, t))$, $\mathbf{I}_2(\mathbf{x}, t) = \mathbf{I}_2(\mathbf{v}_2(\mathbf{x}, t))$ and the

occluding boundary be represented by (4). The frequency spectrum of the occlusion is

$$\begin{aligned} \hat{\mathbf{I}}(\mathbf{k}, \omega) = & \\ & \pi \sum_{\mathbf{n}=-\infty}^{\infty} c_{1\mathbf{n}} \delta(\mathbf{k} - N\mathbf{k}_1, \omega + \mathbf{a}_1^T N\mathbf{k}_1) \\ & + (1 - \pi) \sum_{\mathbf{n}=-\infty}^{\infty} c_{2\mathbf{n}} \delta(\mathbf{k} - N\mathbf{k}_2, \omega + \mathbf{a}_2^T N\mathbf{k}_2) \\ & - i \sum_{\mathbf{n}=-\infty}^{\infty} \left(\frac{c_{1\mathbf{n}} \delta((\mathbf{k} - N\mathbf{k}_1)^T \mathbf{n}_1^-, \mathbf{k}^T \mathbf{a}_1 + \omega)}{(\mathbf{k} - N\mathbf{k}_1)^T \mathbf{n}_1} \right. \\ & \left. + \frac{c_{2\mathbf{n}} \delta((\mathbf{k} - N\mathbf{k}_2)^T \mathbf{n}_1^-, \mathbf{k}^T \mathbf{a}_1 + \omega - \Delta \mathbf{a}^T N\mathbf{k}_2)}{(\mathbf{k} - N\mathbf{k}_2)^T \mathbf{n}_1} \right) \end{aligned} \quad (8)$$

where $\Delta \mathbf{a} = \mathbf{a}_1 - \mathbf{a}_2$. Figure 2 shows the composition of a simple occlusion scene with 2D sinusoidal signals. The Fourier transform of its various components show that the linear spectrum of the occlusion boundary is oriented in a manner that is consonant with the velocity of the occluding signal. Detecting this orientation amounts to identifying the occluding velocity. In addition, since both the occluding signal and its boundary translate with identical full velocity, one can disambiguate the normal velocity of the occluding signal.

4 Generalized Boundaries

Typically, occlusion boundaries are unconstrained in shape, yielding a variety of occluding situations. Under the hypothesis that the motion of the occluding boundary is rigid on the image plane, we can derive the frequency structure of such events. For instance, consider a generalized occlusion boundary represented by the characteristic function $\chi(\mathbf{x})$ in the coordinates

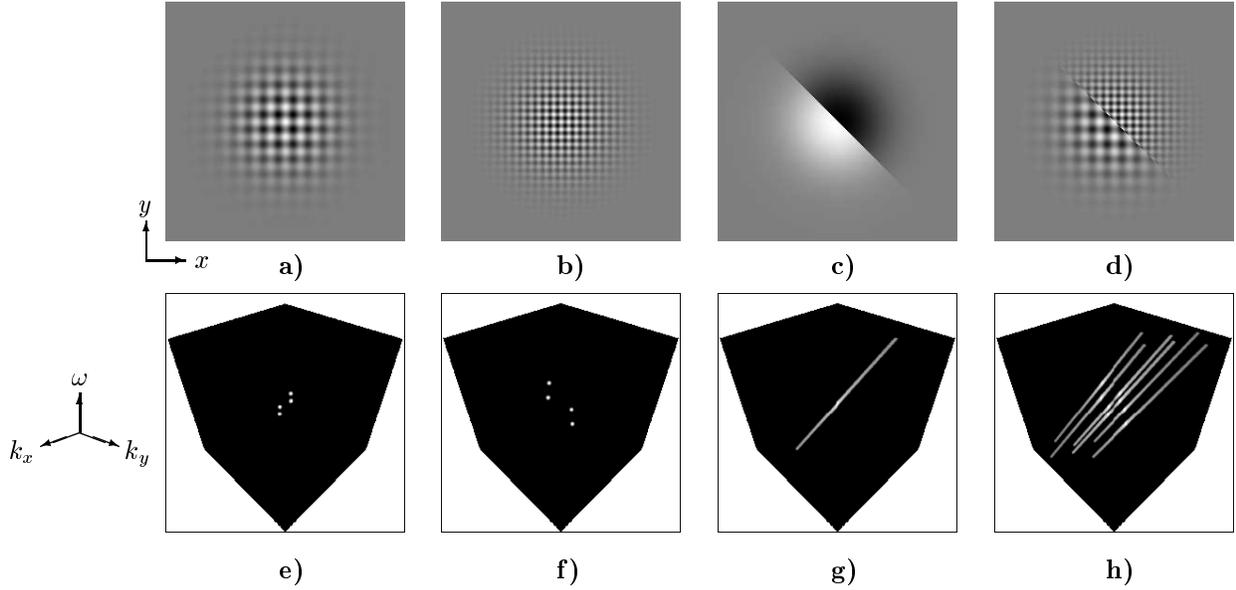


Figure 2: **(top):** The composition of a simple 2D occlusion scene. **a)** The occluding sinusoidal signal with frequency $(\frac{2\pi}{16}, \frac{2\pi}{16})$ and velocity $(-1.0, -1.0)$. **b)** The occluded sinusoidal signal with frequency $(\frac{2\pi}{8}, \frac{2\pi}{8})$ and velocity $(1.0, 1.0)$. **c)** The step function used to create the occlusion scene with normal vector $(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2})$. **d)** The occlusion as a combination of **a)**, **b)** and **c)**. **(bottom) e)** through **h)**: Image plots of corresponding amplitude spectra.

of the image plane and the Fourier transforms of the complex exponential series expansions of both the occluding and occluded signals \mathbf{I}_1 and \mathbf{I}_2 . Substituting these terms into the Fourier transform of (2) yields the following Fourier spectrum

$$\begin{aligned} \hat{\mathbf{I}}(\mathbf{k}, \omega) = & \\ & \sum_{\mathbf{n}=-\infty}^{\infty} c_{1\mathbf{n}} \hat{\chi}(\mathbf{k} - N\mathbf{k}_1) \delta(\mathbf{a}_1^T \mathbf{k} + \omega) - \\ & \sum_{\mathbf{n}=-\infty}^{\infty} c_{2\mathbf{n}} \hat{\chi}(\mathbf{k} - N\mathbf{k}_2) \delta(\mathbf{a}_2^T \mathbf{k} + \omega - \Delta \mathbf{a}^T N\mathbf{k}_2) + \\ & \sum_{\mathbf{n}=-\infty}^{\infty} c_{2\mathbf{n}} \delta(\mathbf{k} - N\mathbf{k}_2, \omega + \mathbf{a}_2^T N\mathbf{k}_2), \end{aligned} \quad (9)$$

from which it is observed that the spectrum of the occluding boundary is repeated at every non-zero frequency of both signals. The spectrum occupies planes descriptive of full velocity which can be used to perform such measurements.

5 Theta Motion

Sources of Non-Fourier motion include such phenomena as translucency, occlusion and, in particular, Zanker's

Theta motion stimuli which are examples of Non-Fourier motion involving occlusion with various forms of occlusion windows[5]. This category of motion is described by a window that translates with a velocity uncorrelated with those of the occluding and occluded signals. For 1D image signals, such an occlusion scene can be expressed as

$$\begin{aligned} \mathbf{I}(x, t) = & \chi(x - v_3 t) \mathbf{I}_1(x - v_1 t) \\ & - \chi(x - v_3 t) \mathbf{I}_2(x - v_2 t) \\ & + \mathbf{I}_2(x - v_2 t). \end{aligned} \quad (10)$$

As Zanker and Fleet [5, 2], we model the occlusion window with a rectangle function in the spatial coordinate as

$$\chi\left(\frac{x - x_0}{b}\right) = \begin{cases} 0 & \text{if } \left|\frac{x - x_0}{b}\right| > \frac{1}{2} \\ \frac{1}{2} & \text{if } \left|\frac{x - x_0}{b}\right| = \frac{1}{2} \\ 1 & \text{if } \left|\frac{x - x_0}{b}\right| < \frac{1}{2}. \end{cases} \quad (11)$$

Such a function is non-zero in the interval $]x_0 - \frac{b}{2}, x_0 + \frac{b}{2}[$ and zero otherwise. We can then write the Fourier transform of the occlusion scene (10) as

$$\begin{aligned} \mathbf{I}(k, \omega) = & \\ & K \sum_{\mathbf{n}=-\infty}^{\infty} \text{sinc}(k - n\mathbf{k}_1) c_{1\mathbf{n}} \delta(k v_3 - n\mathbf{k}_1 \Delta v_3) - \end{aligned}$$

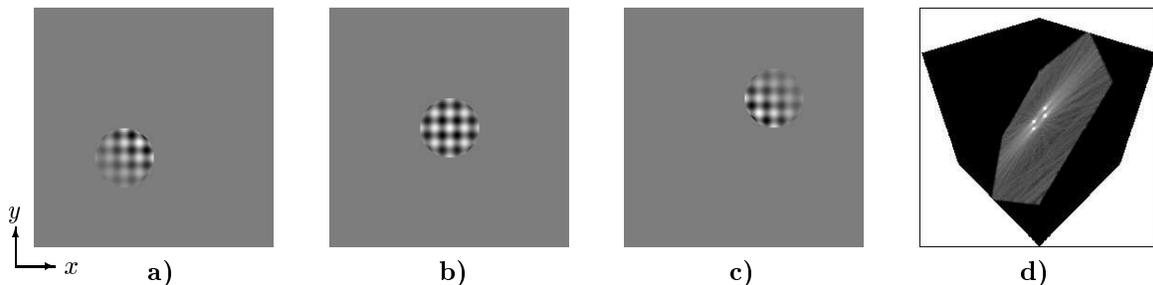


Figure 3: *Generalized occluding boundaries. a), b) and c): Images from a sequence in which the occluding pattern moves with velocity $\mathbf{a}_1^T = (-1.0, -1.0)$. Spatial frequency of the sinusoidal texture within the circular boundary is $\mathbf{k}_1^T = (\frac{2\pi}{16}, \frac{2\pi}{16})$. d): The frequency spectrum of the sequence, where the plane contains the spectrum of the boundary convolved with the frequency of the texture.*

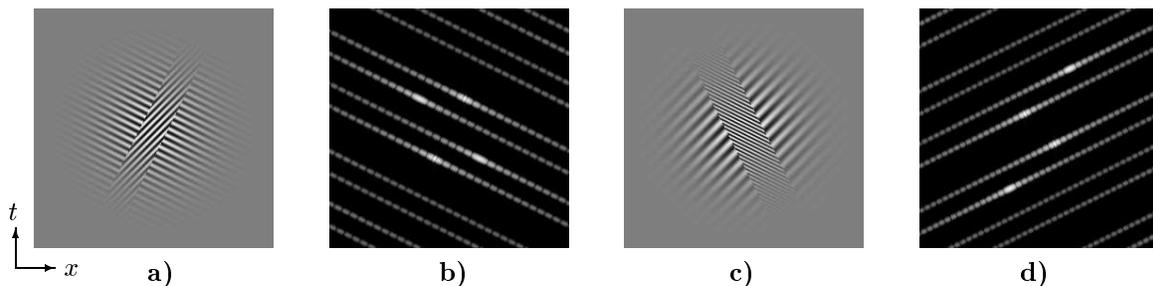


Figure 4: *Examples of Theta motion. a): Velocities of occlusion window, occluding and occluded signals are $v_3 = 0.5$, $v_1 = 1.0$ and $v_2 = -2.0$ respectively. b): Frequency spectrum of a). c): Velocities of occlusion window, occluding and occluded signals are $v_3 = -0.5$, $v_1 = -2.0$ and $v_2 = 1.0$ respectively. d): Frequency spectrum of c).*

$$K \sum_{n=-\infty}^{\infty} \text{sinc}(k - nk_2) c_{2n} \delta(kv_3 - nk_2 \Delta v_2) + \sum_{n=-\infty}^{\infty} c_{2n} \delta(k - nk_2, \omega + nk_2 v_2), \quad (12)$$

where $\text{sinc}(k) = \frac{\sin k}{k}$, $\Delta v_3 = v_3 - v_1$, $\Delta v_2 = v_2 - v_1$, and $K = b^{-1} e^{-ikx_0 b^{-1}}$. The spectra $\delta(kv_3 + \omega - nk_1 \Delta v_3)$ and $\delta(kv_3 + \omega - nk_2 \Delta v_2)$ are consonant with the motion of the occluding window and represent a case of Non-Fourier motion, as they do not contain the origin.

We performed two experiments with Theta motions as pictured in Figure 4. It is observed that the spectrum of the sinc function is convolved with each frequency of both signals and that its orientation is descriptive of the velocity of the window. As expected, the visible peaks represent the motions of both signals in the MFFC sense.

6 Conclusion

We presented a framework for the Fourier analysis of non-Fourier motions which yields exact mathematical expressions and forms a basis for computational models for velocity extraction in the frequency domain beyond the MFFC principle.

References

- [1] C. Chubb and G. Sperling. Drift-balanced random stimuli: A general basis for studying non-fourier motion perception. *J. Opt. Soc. Am. A*, 5(11):1986–2007, 1988.
- [2] D. J. Fleet and K. Langley. Computational analysis of non-fourier motion. *Vision Research*, 34(22):3057–3079, 1995.

- [3] K. Turano and A. Pantle. On the mechanism that encodes the movement of contrast variations. *Vision Research*, 29:207–221, 1989.
- [4] J. D. Victor and M. M. Conte. Coherence and transparency of moving plaids composed of fourier and non-fourier gratings. *Perception & Psychophysics*, 52(4):403–414, 1988.
- [5] M. J. Zanker. Theta motion: A paradoxical stimulus to explore higher-order motion extraction. *Vision Research*, 33:553–569, 1993.