A theory of occlusion in the context of optical flow

Steven S. Beauchemin and John L. Barron

1 Introduction

Traditionally, image motion and its approximation known as optical flow have been treated as continuous functions of the image domain [9]. However, in realistic imagery, one finds cases verifying this hypothesis exceedingly rarely. Many phenomena may cause discontinuities in the optical flow function of imagery [16]. Among them, occlusion and translucency are frequent causes of discontinuities in realistic imagery. In addition, their information content is useful to later stages of processing [8] such as motion segmentation [1] and 3-d surface reconstruction [17].

Occlusion boundaries are described as the partial occlusion of a surface by another, while translucency is defined as occlusion of a surface by translucent material. In realistic imagery, one finds occlusion to be the most frequent cause of discontinuous motion.

Recently, a number of algorithms have been designed to handle multiple motions [2]. Among these, we find constraint-line clustering algorithms [15], which use a form of cluster analysis applied to sets of constraint lines to determine the dominant motion of a given image region. Similarly, robust estimators are used to recover dominant motion [6]. However, approaches focusing on the determination of dominant motion do not explicitly form a multiple motion model in the sense that they only provide one velocity measurement in regions where many motions may prevail. Alternatively, a number of authors have studied inhibitory smoothness constraints [13] which relax smoothness requirements at image regions of high grayvalue gradients. Nonetheless, intensity discontinuities may not necessarily represent motion discontinuities. Other approaches consist of refining the boundaries of closed curves delimiting regions exhibiting coherent motion [14]. Again, such schemes are limited by an explicit single motion model.

Schemes for estimating optical flow at regions of partial translucency have also appeared in the recent literature [5]. In the case of spatiotemporal frequency models [8], translucency could be easily handled, given an appropriate constraint-integration algorithm. In addition, schemes capable of handling both occlusion and translucency have been devised. For instance, sequential tracking and registration algorithms which independently compute the velocity of objects from a scene exhibiting different motions, while handling both occlusion and translucency, have been used [10]. Multilayer and superposition models also provide paradigms for the estimation of multiple velocities. These approaches consider the optical flow function of the image as a superposition of motion layers, each one described by a unique set of parameters, thus allowing discontinuous and multiple-valued optical flow functions [7, 11, 16].

These approaches, while constituting valuable contributions, do not provide descriptive models of phenomena such as occlusion and translucency. Further, it is often difficult if not impossible to determine the nature of the image events giving rise to multiple motions. In these schemes, occlusion may not be differentiated from translucency and the explicit identification of the motions associated with both the occluding and occluded signals remains an open problem. In this contribution we demonstrate, under a minimal set of assumptions, that such distinctions can be made through a Fourier analysis of these image events. We also show that translucency may be handled as a special case of occlusion.

2 Multiple motions

Given an arbitrary environment and a moving visual sensor, the motion field generated onto the imaging plane by a 3-d scene within the visual field is represented as a function of the motion parameters of the visual sensor, usually expressed as instantaneous translation $\mathbf{T}^T = (T_x, T_y, T_z)$ and rotation $\mathbf{\Omega}^T = (\Omega_x, \Omega_y, \Omega_z)$:

$$\mathbf{v} = \begin{pmatrix} Z(\mathbf{x})^{-1}(xT_z - T_x) + xy\Omega_x - (1+x^2)\Omega_y + y\Omega_z \\ Z(\mathbf{x})^{-1}(yT_z - T_y) + (1+y^2)\Omega_x - xy\Omega_y - x\Omega_z \end{pmatrix}, \tag{1}$$

where $\mathbf{x}^T = (x, y)$ is the perspective projection of a point $\mathbf{P}^T = (X, Y, Z)$ in the visual field. Assuming that the motion of the visual sensor is continuous (that is to say: $\mathbf{\Omega}$ and $\mathbf{T}$ are differentiable with respect to time), discontinuities in image motion are then introduced in (1) whenever the depth function $Z(\mathbf{x})$ is other than single-valued and differentiable. The occurrence of occlusion causes the depth function to exhibit a discontinuity, whereas translucency leads to a multiple-valued depth function.
Our motivation for conducting this study is to use occlusion phenomena as a source of information rather than an obstacle in conflict with too simple hypotheses concerning the structure of optical flow fields, such as the single surface assumption. Crucial information, such as the identification of image events leading to multiple motions, the determination of these motions and, under occlusion, their identification as velocities of occluding and occluded surfaces, can be obtained through a Fourier analysis of occlusion and transparency phenomena.
Throughout this contribution, we consider velocities as locally constant quantities and image signals as satisfying Dirichlet conditions.

2.1 Occlusion in the frequency domain

The Fourier transform of the optical flow constraint equation is obtained with the differentiation property as:

$$\mathcal{F}\left[\nabla \mathbf{I}(\mathbf{x}, t)^T \mathbf{v} + \mathbf{I}_t\right] = i\hat{\mathbf{I}}(\mathbf{k}, \omega)\delta(\mathbf{k}^T\mathbf{v} + \omega), \tag{2}$$

where i is the imaginary number, $\hat{\mathbf{I}}(k, \omega)$ is the Fourier transform of $\mathbf{I}(x, t)$ and $\delta(\mathbf{k}^T\mathbf{v} + \omega)$ is a line-mass Dirac delta function. Expression (2) yields $\mathbf{k}^T\mathbf{v} + \omega = 0$

as a constraint on velocity. Similarly, the Fourier transform of a translating intensity profile $\mathbf{I}(\mathbf{x},t)$ is obtained with the shift property as:

$$\begin{aligned}\hat{\mathbf{I}}(\mathbf{k},\omega) &= \int\int \mathbf{I}(\mathbf{x}-\mathbf{v}t)e^{-i(\mathbf{k}^T\mathbf{x}+\omega t)}d\mathbf{x}dt \\ &= \hat{\mathbf{I}}(\mathbf{k})\delta(\mathbf{k}^T\mathbf{v}+\omega), \end{aligned} \tag{3}$$

which also yields the constraint $\mathbf{k}^T\mathbf{v}+\omega=0$. Hence, (2) and (3) demonstrate that the frequency analysis of image motion is in accordance with the motion constraint equation [8]. It is also observed that $\mathbf{k}^T\mathbf{v}+\omega=0$ represents, in the frequency domain, an oriented plane passing through the origin, with normal vector $\mathbf{v}$ representing full velocity, onto which the Fourier spectrum of $\mathbf{I}(\mathbf{x})$ lies[1]. The discontinuities in optical flow arising from occlusion may be written by considering two translating intensity profiles, one partially occluding the other. Let $\mathbf{I}_1(\mathbf{x})$ and $\mathbf{I}_2(\mathbf{x})$ be the intensity profiles of an object and a background scene. An object indicator function such as

$$\mathbf{U}(\mathbf{x}) = \begin{cases} 1 & \text{if } \mathbf{I}_1(\mathbf{x}) \neq 0 \\ 0 & \text{otherwise} \end{cases}$$

may be defined to specify the actual location of the object on the image plane. The resulting intensity pattern is then written as a function of the intensity profiles of the object, the background and the object indicator:

$$\mathbf{I}(\mathbf{x},t) = \mathbf{I}_1(\mathbf{x}-\mathbf{v}_1t) + \left[1-\mathbf{U}(\mathbf{x}-\mathbf{v}_1t)\right]\mathbf{I}_2(\mathbf{x}-\mathbf{v}_2t). \tag{4}$$

By using the shift property of Fourier transforms, (4) is rewritten in spatiotemporal frequency space as:

$$\begin{aligned}\hat{\mathbf{I}}(\mathbf{k},\omega) &= \hat{\mathbf{I}}_1(\mathbf{k})\delta(\omega+\mathbf{v}_1^T\mathbf{k}) + \hat{\mathbf{I}}_2(\mathbf{k})\delta(\omega+\mathbf{v}_2^T\mathbf{k}) \\ &- \left[\hat{\mathbf{U}}(\mathbf{k})\delta(\omega+\mathbf{v}_1^T\mathbf{k})\right] * \left[\hat{\mathbf{I}}_2(\mathbf{k})\delta(\omega+\mathbf{v}_2^T\mathbf{k})\right]. \end{aligned} \tag{5}$$

The first two terms of (5) are the signals associated with the object and the background. The frequency spectra of $\mathbf{I}_1$ and $\mathbf{I}_2$ are located on the planes defined by the equations $\mathbf{k}^T\mathbf{v}_1+\omega=0$ and $\mathbf{k}^T\mathbf{v}_2+\omega=0$ respectively. In addition, the respective orientations of these planes fully determine $\mathbf{v}_1$ and $\mathbf{v}_2$. The last term of (5) describes the distortion created by the occlusion boundary. In the following sections, this form of distortion is analized and its usefulness in determining image events giving rise to multiple motions is shown.

[1] The aperture problem arises when the Fourier spectrum of $\mathbf{I}(\mathbf{x})$ is concentrated on a line rather than on a plane [8, 12]. Spatiotemporally, this depicts the situation in which $\mathbf{I}(\mathbf{x})$ exhibits a single orientation. In this case, one may only obtain the speed and direction of motion normal to the orientation, noted as $\mathbf{v}_\perp$. If many normal velocities are found in a single neighbourhood, their respective lines fit the plane $\mathbf{k}^T\mathbf{v}+\omega=0$ from which full velocity may be obtained.

3 Frequency analysis of occlusion

The analysis begins with the consideration of a simple case consisting of two 1-d sinusoidal intensity profiles. The results are then generalized to arbitrary 1-d and 2-d intensity profiles.

3.1 One-dimensional sinusoidal signals

The case in which two 1-d sinusoidals play the role of the object and the background is first considered. Let $\mathbf{I}(x,t)$ be a 1-d intensity function $\mathbf{I}_1(x)$ translating with velocity v_1: $\mathbf{I}(x,t) = \mathbf{I}_1(x - v_1 t)$. Its Fourier transform is $\hat{\mathbf{I}}(k,\omega) = \hat{\mathbf{I}}_1(k)\delta(kv_1 + \omega)$. Let $\mathbf{I}_1(x)$ be occluding another 1-d intensity pattern $\mathbf{I}_2(x)$ moving with velocity v_2. The resulting intensity profile can then be expressed as:

$$\mathbf{I}(x,t) = \mathbf{u}(x - v_1 t)\mathbf{I}_1(x - v_1 t) + (1 - \mathbf{u}(x - v_1 t))\mathbf{I}_2(x - v_2 t) \tag{6}$$

where $\mathbf{u}(x)$ is Heaviside's function representing the occluding point:

$$\mathbf{u}(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ 0 & \text{otherwise.} \end{cases}$$

The Fourier transform of the intensity profile (6) is:

$$\begin{aligned} \hat{\mathbf{I}}(k,\omega) &= [\hat{\mathbf{u}}(k)\delta(kv_1 + \omega)] * [\hat{\mathbf{I}}_1(k)\delta(kv_1 + \omega)] \\ &- [\hat{\mathbf{u}}(k)\delta(kv_1 + \omega)] * [\hat{\mathbf{I}}_2(k)\delta(kv_2 + \omega)] \\ &+ \hat{\mathbf{I}}_2(k)\delta(kv_2 + \omega), \end{aligned} \tag{7}$$

where $\hat{\mathbf{u}}(k)$ is the Fourier transform of Heaviside's function $\mathbf{u}(x)$ written as $\hat{\mathbf{u}}(k) = \pi\delta(k) + (ik)^{-1}$.

Proposition 1 *Let $\mathbf{I}_1(x)$ and $\mathbf{I}_2(x)$ be cosine functions with respective angular frequencies $k_1 = 2\pi f_1 > 0$ and $k_2 = 2\pi f_2 > 0$ and let $\mathbf{I}_1(x - v_1 t) = c_1 \cos(k_1 x - v_1 t)$ and $\mathbf{I}_2(x - v_2 t) = c_2 \cos(k_2 x - v_2 t)$. Then the frequency spectrum of the occlusion obtained by substituting $\mathbf{I}_1(x)$ and $\mathbf{I}_2(x)$ into (6) is:*

$$\begin{aligned} \hat{\mathbf{I}}(k,\omega) &= \frac{\pi}{2} c_1 \delta(k \pm k_1, \omega \mp k_1 v_1) \\ &+ \frac{(1-\pi)}{2} c_2 \delta(k \pm k_2, \omega \mp k_2 v_2) \\ &+ \frac{i}{2}\left(\frac{c_2 \delta(kv_1 + \omega \pm k_2 \Delta v)}{(k \pm k_2)} - \frac{c_1 \delta(kv_1 + \omega)}{(k \pm k_1)} \right) \end{aligned} \tag{8}$$

A number of conclusions can be drawn from proposition 1: Since the signals are cosines, all their power content is real. In addition, the power content of the distortion term is entirely imaginary, and form lines of decreasing power about the frequencies of both the occluding and occluded signals. Their orientation is inversely proportional to the velocity of the occluding signal, as $-v_1$ is the slope of the constraint lines.

3.2 One-dimensional arbitrary signals

In general, the occluding and occluded signals cannot be represented as simple sinusoidal functions. To gain generality, $\mathbf{I}_1(x)$ and $\mathbf{I}_2(x)$ may be expanded as a series of complex exponentials, assuming that functions $\mathbf{I}_1(x)$ and $\mathbf{I}_2(x)$ satisfy Dirichlet conditions.

Proposition 2 *Let $\mathbf{I}_1(x)$ and $\mathbf{I}_2(x)$ be functions satisfying Dirichlet conditions such that they may be expressed as complex exponential series expansions:*

$$\mathbf{I}_1(x) = \sum_{n=-\infty}^{\infty} c_{1n} e^{ink_1 x} \quad \mathbf{I}_2(x) = \sum_{n=-\infty}^{\infty} c_{2n} e^{ink_2 x}, \tag{9}$$

where n is integer, c_{1n} and c_{2n} are complex coefficients and k_1 and k_2 are the fundamental frequencies of both signals. Then the frequency spectrum of the occlusion obtained by substituting the frequency spectra of (9) into (6) is:

$$\begin{aligned}
\hat{\mathbf{I}}(k,\omega) &= \pi \sum_{n=\infty}^{\infty} c_{1n}\delta(k - nk_1, \omega + nk_1 v_1) \\
&+ (1-\pi) \sum_{n=-\infty}^{\infty} c_{2n}\delta(k - nk_2, \omega + nk_2 v_2) \\
&+ i \sum_{n=-\infty}^{\infty} \left(\frac{c_{2n}\delta(kv_1 + \omega - nk_2\Delta v)}{(k - nk_2)} - \frac{c_{1n}\delta(kv_1 + \omega)}{(k - nk_1)} \right).
\end{aligned} \tag{10}$$

Proposition 2 is an important generalization of the first one: Any signal which represents a physical quantity satisfies Dirichlet conditions and therefore may be expressed as an expansion of complex exponentials. Since c_{1n} and c_{2n} are complex coefficients, the power contents of the signals are both real and imaginary.

3.3 Two-dimensional arbitrary signals

Imagery is the result of the projection of light reflected by environmental features onto the imaging plane of the visual sensor. Hence, such signals are inherently two dimensional. Towards a generalization of (10), (9) is expanded as series of 2-d complex exponentials.

Proposition 3 *Let $\mathbf{I}_1(\mathbf{x})$ and $\mathbf{I}_2(\mathbf{x})$ be 2-d functions satisfying Dirichlet conditions such that they may be expressed as complex exponential series expansions:*

$$\mathbf{I}_1(\mathbf{x}) = \sum_{\mathbf{n}=-(\infty,\infty)}^{(\infty,\infty)} c_{1\mathbf{n}} e^{i\mathbf{x}^T N \mathbf{k}_1} \quad \mathbf{I}_2(\mathbf{x}) = \sum_{\mathbf{n}=-(\infty,\infty)}^{(\infty,\infty)} c_{2\mathbf{n}} e^{i\mathbf{x}^T N \mathbf{k}_2}, \tag{11}$$

where $\mathbf{n}^T = (n_x, n_y)$ are integers, $N = \mathbf{n}^T I$, $\mathbf{x}^T = (x, y)$ are spatial coordinates, $\mathbf{k}_1 = (k_{x1}, k_{y1})$ and $\mathbf{k}_2 = (k_{x2}, k_{y2})$ are spatial frequencies and $c_{1\mathbf{n}}$ and $c_{2\mathbf{n}}$ are complex coefficients. Also let the occluding boundary be locally represented by:

$$\mathbf{U}(\mathbf{x}) = \begin{cases} 1 & \text{if } \mathbf{x}^T \vec{\eta} \geq 0 \\ 0 & \text{otherwise,} \end{cases} \tag{12}$$

where $\vec{\eta}$ is a vector normal to the instantaneous slope of the occluding boundary at $\mathbf{x}$. Then the frequency spectrum of the occlusion obtained by substituting the frequency spectra of (11) and (12) into a 2-d version of (6) is:

$$\begin{aligned}\hat{\mathbf{I}}(\mathbf{k},\omega) &= \pi \sum_{\mathbf{n}=-(\infty,\infty)}^{(\infty,\infty)} c_1\mathbf{n}\delta(\mathbf{k}-N\mathbf{k}_1,\omega+\mathbf{v}_1^T N\mathbf{k}_1) \\ &+ (1-\pi) \sum_{\mathbf{n}=-(\infty,\infty)}^{(\infty,\infty)} c_2\mathbf{n}\delta(\mathbf{k}-N\mathbf{k}_2,\omega+\mathbf{v}_2^T N\mathbf{k}_2) \\ &+ i \sum_{\mathbf{n}=-(\infty,\infty)}^{(\infty,\infty)} \left(\frac{c_2\mathbf{n}\delta(\mathbf{k}^T\mathbf{v}_1+\omega-\Delta\mathbf{v}^T N\mathbf{k}_2)}{(\mathbf{k}-N\mathbf{k}_2)^T\vec{\eta}} - \frac{c_1\mathbf{n}\delta(\mathbf{k}^T\mathbf{v}_1+\omega)}{(\mathbf{k}-N\mathbf{k}_1)^T\vec{\eta}} \right), \end{aligned} \tag{13}$$

where $\mathbf{v}_1^T=(u_1,v_1)$, $\mathbf{v}_2^T=(u_2,v_2)$ and $\Delta\mathbf{v}=\mathbf{v}_1-\mathbf{v}_2$.

Proposition 3 is a direct extension of proposition 2 in two spatial dimensions. For this general case, the constraint lines of propositions 1 and 2 generated by both the occluding and occluded signals and the distortion terms become constraint planes. The frequency structures of individual signals are preserved to within scaling factors.

3.4 Relation to translucency

Transmission of light through translucent material may cause multiple motions to arise in the same image region. Generally, this effect is depicted on the image plane as

$$\mathbf{I}(\mathbf{x},t) = f(\rho_1)(\mathbf{x}-\mathbf{v}_1 t)\mathbf{I}_2(\mathbf{x}-\mathbf{v}_2 t), \tag{14}$$

where $f(\rho_1)$ is a function of the density of the translucent material [8]. Under the local assumption of spatially constant $f(\rho_1)$, with translucency factor φ, (14) is reformulated as a weighted superposition of intensity profiles:

$$\mathbf{I}(\mathbf{x},t) = \varphi\mathbf{I}_1(\mathbf{x}-\mathbf{v}_1 t) + (1-\varphi)\mathbf{I}_2(\mathbf{x}-\mathbf{v}_2 t), \tag{15}$$

where $\mathbf{I}_1(\mathbf{x},t)$ is the intensity profile of the translucent material and $\mathbf{I}_2(\mathbf{x},t)$ is the intensity profile of the background. With $\mathbf{I}_1(\mathbf{x})$ and $\mathbf{I}_2(\mathbf{x})$ satisfying Dirichlet conditions, the frequency spectrum of (15) is written as:

$$\begin{aligned}\hat{\mathbf{I}}(\mathbf{k},\omega) &= \varphi \sum_{\mathbf{n}=-(\infty,\infty)}^{(\infty,\infty)} c_1\mathbf{n}\delta(\mathbf{k}-N\mathbf{k}_1,\omega+\mathbf{v}_1^T N\mathbf{k}_1) \\ &+ (1-\varphi) \sum_{\mathbf{n}=-(\infty,\infty)}^{(\infty,\infty)} c_2\mathbf{n}\delta(\mathbf{k}-N\mathbf{k}_2,\omega+\mathbf{v}_2^T N\mathbf{k}_2). \end{aligned} \tag{16}$$

With the exception of the distortion term, and to within scaling factors, (16) is identical to (13). Hence, with respect to its frequency structure, translucency may be reduced to a special case of occlusion for which the distortion terms vanish.

3.5 Geometric interpretation

In the simplest case involving sinusoidal signals, Proposition 1 shows that the frequency spectra of both signals are preserved to within scaling factors. In addition, the imaginary terms represent the frequency spectrum of the occlusion boundary. Figure 2 shows one case of occlusion with a 1-d, Gaussian-windowed sinusoidal signal. The velocity of the occluding signal is $v_1 = 1.0$. The velocity of the occluded signal is $v_2 = -1.0$. The spatial frequency of the occluding and occluded signals are $k_1 = \frac{2\pi}{16}$ and $k_2 = \frac{2\pi}{8}$ respectively. The vertical axis represents temporal frequency ω while the horizontal axis is spatial frequency k. The spectral peaks located at $\pm(k_1, -k_1 v_1)$ and $\pm(k_2, -k_2 v_2)$ depict the spatiotemporal frequencies of both signals and fit the constraint lines $kv_1 + \omega = 0$ and $kv_2 + \omega = 0$. The oblique spectra intesecting the peaks represent the spectrum generated by the occlusion boundary and fit the constraint lines $k_1 v_1 + \omega \pm k_2 v_2 = 0$ and $k_1 v_1 + \omega = 0$. These lines are parallel to the constraint line of the occluding signal. It also is interesting to observe From Proposition 2 that every non-zero frequency of an occluded signal shows such a parallel line due to occlusion.

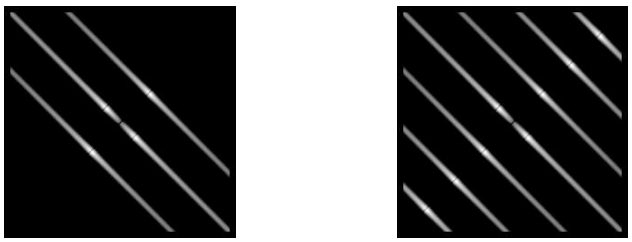

Figure 1: **a) (left):** Gaussian-smoothed frequency spectrum produced with (8). The occluding and occluded signals have frequencies $k_1 = \frac{2\pi}{16}$ and $k_2 = \frac{2\pi}{8}$ and velocities $v_1 = 1$ and $v_2 = -1$ respectively. **b) (right):** Gaussian-smoothed frequency spectrum produced with (10). The occluding signal has frequency $k_1 = \frac{2\pi}{16}$ and velocity $v_1 = 1$ while the occluded signal is composed of frequencies $nk_2 = 2\pi\left(\frac{n}{8}\right), n = 1, 2, 3$ and velocity $v_2 = -1$.

Proposition 3 is the generalization of Proposition 2 in 2D and its geometric interpretation is similar. That is to say, the constraint lines of the signals and the occlusion boundary become constraint planes. For instance, the frequencies $(N\mathbf{k}_1, -\mathbf{v}_1^T N\mathbf{k}_1)$ and $(N\mathbf{k}_2, -\mathbf{v}_2^T N\mathbf{k}_2)$ fit the constraint planes of the occluding and occluded signals, defined as $\mathbf{k}_1^T \mathbf{v}_1 + \omega = 0$ and $\mathbf{k}^T \mathbf{v}_2 = 0$. In the distortion term, the arguments of the Dirac δ functions $\mathbf{k}^T \mathbf{v}_1 + \omega - \Delta\mathbf{v}^T N\mathbf{k}_2$ and $\mathbf{k}^T \mathbf{v}_1 + \omega$ represent a set of planes parallel to the constraint plane of the occluding signal $\mathbf{k}^T \mathbf{v}_1 + \omega = 0$. That is to say, for every discrete frequency $N\mathbf{k}_1$ and $N\mathbf{k}_2$ exhibited by both signals, there is a frequency spectrum fitting the planes given by $\mathbf{k}^T \mathbf{v}_1 + \omega - \Delta\mathbf{v}^T N\mathbf{k}_2 = 0$ and $\mathbf{k}^T \mathbf{v}_1 + \omega = 0$. The magnitudes of these planar spectra are determined by their corresponding scaling functions $c_1 \mathbf{n}[(\mathbf{k} - N\mathbf{k}_1)^T \vec{\eta}\,]^{-1}$ and $c_2 \mathbf{n}[(\mathbf{k} - N\mathbf{k}_2)^T \vec{\eta}\,]^{-1}$. Hence, Proposition 3 reveals useful constraint planes, as the

power spectra of both signals peak within planes $\mathbf{k}^T\mathbf{v}_1+\omega=0$ and $\mathbf{k}^T\mathbf{v}_2+\omega=0$ and the constraint planes arising from the distortion are parallel to the spectrum of the occluding signal $\mathbf{I}_1(\mathbf{x},t)$.

4 Numerical experiments

Aside from deriving formal proofs[2], several experiments were performed in support of the Propositions. The Fourier spectra obtained with both a standard FFT algorithm and those predicted by the theory were compared.
In order to verify the propositions, two 1-d signals which respectively act as occluding and occluded surfaces were used. Expression (6) is used with $\mathbf{I}_1(x-v_1t)=c_1\cos(k_1x-v_1t)$ and $\mathbf{I}_2(x-v_2t)=c_2\cos(k_2x-v_2t)$, where $\mathbf{I}_1$ and $\mathbf{I}_2$ are the occluding and the occluded surfaces with respective frequencies $k_1=\frac{2\pi}{16}$ and $k_2=\frac{2\pi}{8}$. Constants c_1 and c_2 correspond to signal amplitudes. To limit boundary conditions when numerically computing Fourier transforms, the signal was windowed with a Gaussian envelope. The discrete Fourier transform of the

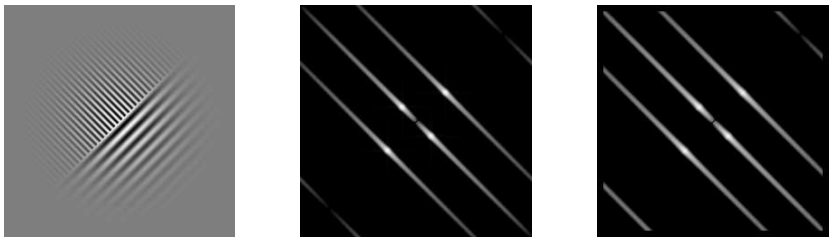

Figure 2: **a) (left):** Gaussian-windowed signal with sinusoidals acting as occluding and occluded surfaces. The occluding signal has frequency $k_1=\frac{2\pi}{16}$ and velocity $v_1=1$. The occluded signal has frequency $k_2=\frac{2\pi}{8}$ and velocity $v_2=-1$. **b) (middle):** Fourier spectrum generated with a standard FFT algorithm. **c) (right):** Fourier spectrum predicted by theory.

windowed signal obtained with a standard FFT algorithm is shown in Figure 2b, where the peaks associated with both sinusoidals and the distortion lines are clearly visible. A discretized version of proposition 1, which models aliasing effects is shown in Figure 2c for the same frequencies and velocities as in Figure 2b: The spectra obtained with both a standard FFT algorithm and the theoretical results are essentially identical.

[2] The formal proofs of Propositions 1, 2 and 3 are found in [3]

5 Conclusion

Under a minimal set of hypotheses, such as locally constant velocity and intensity profiles satisfying Dirichlet conditions, we have shown the Fourier structure of occlusion and translucency phenomena in both 1 and 2D and outlined various interesting geometrical properties. For instance, the constraint lines or planes cast by the occlusion boundary have been characterized: In a multiple motion situation, their presence indicates an occlusion while their absence indicates a translucency phenomenon.
When a multiple motion situation is caused by an occlusion, the parallelism between the the distortion cast by the occluding boundary and the Fourier spectrum of the occluding signal differentiates the velocity of the occluding signal from the velocity of the occluded signal.
This analysis forms a basis for a class of constraint-grouping algorithms capable of distinguishing occlusion from translucency and identifying occluded and occluding surfaces, along with their respective velocities. In addition, the inclusion of the occlusion boundary distortions in models of multiple motions is likely to result in further improvements of signal-to-noise ratios. The results of this theory have been extended to linear models of optical flow and signal degeneracy caused by the aperture problem [4].

References

[1] G. Adiv. Determining three-dimensional motion and structure from optical flow generated by several moving objects. *IEEE PAMI*, 7(4):384–401, 1985.

[2] S. S. Beauchemin and J. L. Barron. The computation of optical flow. *ACM Computing Surveys*, 27(3):433–467, 1995.

[3] S. S. Beauchemin and J. L. Barron. A theory of occlusion. Technical Report TR-449, Dept. of Computer Science, Univ. of Western Ontario, March 1995.

[4] S. S. Beauchemin, A. Chalifour, and J. L. Barron. Discontinuous optical flow: Recent theoretical results. In *Vision Interface*, pages 57–64, Kelowna, Canada, May 1997.

[5] J. R. Bergen, P. J. Burt, R. Hingorani, and S. Peleg. Three-frame algorithm for estimating two-component image motion. *IEEE PAMI*, 14(9):886–896, 1992.

[6] M. J. Black and P. Anandan. A model for the detection of motion over time. In *Proceedings of ICCV*, pages 33–37, Osaka, Japan, December 1990.

[7] T. Darrell and A. Pentland. Robust estimation of a multi-layered motion representation. In *IEEE Proceedings of Workshop on Visual Motion*, pages 173–178, Princeton, New Jersey, October 1991.

[8] D. J. Fleet. *Measurement of Image Velocity*. Kluwer Academic Publishers, Norwell, 1992.

[9] B. K. P. Horn and B. G. Schunck. Determining optical flow. *Artificial Intelligence*, 17:185–204, 1981.

[10] M. Irani, B. Rousso, and S. Peleg. Computing occluding and transparent motions. *IJCV*, 12(1):5–16, 1994.

[11] A. D. Jepson and M. Black. Mixture models for optical flow computation. In *IEEE Proceedings of CVPR*, pages 760–761, New York, New York, June 1993.

[12] D. Marr and S. Ullman. Directional selectivity and its use in early visual processing. *Proceedings of Royal Society London*, B 211:151–180, 1981.

[13] H.-H. Nagel. Displacement vectors derived from second-order intensity variations in image sequences. *CVGIP*, 21:85–117, 1983.

[14] C. Schnorr. Computation of discontinuous optical flow by domain decomposition. *IEEE PAMI*, 8(2):153–165, 1992.

[15] B. G. Schunck. Image flow segmentation and estimation by constraint line clustering. *IEEE PAMI*, 11(10):1010–1027, 1989.

[16] M. Shizawa and K. Mase. Principle of superposition: A common computational framework for analysis of multiple motion. In *IEEE Proceedings of Workshop on Visual Motion*, pages 164–172, Princeton, New Jersey, October 1991.

[17] P. Toh and A. K. Forrest. Occlusion detection in early vision. In *ICCV*, pages 126–132. IEEE, 1990.

Steven S. Beauchemin and John L. Barron
beau@csd.uwo.ca barron@csd.uwo.ca
The University of Western Ontario
London, Canada
N6A 5B7