The Structure of Occlusion in Fourier Space

S. S. Beauchemin
Dept. of Computer Science
The University of Western Ontario
London, Canada N6A 5B7
e-mail: *beau@csd.uwo.ca*

J. L. Barron
Dept. of Computer Science
The University of Western Ontario
London, Canada N6A 5B7
e-mail: *barron@csd.uwo.ca*

Abstract

A fundamental problem in processing sequences of images is the computation of *optic flow*, an approximation of the motion field which is the projection of velocities of 3-d surface points onto the imaging plane of a visual sensor. A wide variety of algorithms for computing optical flow has been reported in the literature. However, it is not until recently that multiple motion paradigms were seriously studied. Such approaches are necessary to encompass 2-d image events such as occlusion and translucency which violate the classical single motion hypothesis. While current approaches to this problem constitute valuable contributions, they often lack the ability to determine the nature of the image events giving rise to multiple motions: Occlusion may not be differentiated from translucency and it remains difficult to explicitly identify the motions associated with both the occluding and occluded surfaces. We demonstrate, under a set of reasonable 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.

1 Introduction

Many phenomena can cause multiple image motions [20]. Among them, occlusion and translucency are important in terms of their occurrence and significance in realistic imagery. In addition, their information content is useful to later stages of processing [9] such as motion segmentation [1] and 3-d surface reconstruction [21]. Occlusion boundaries are described by 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 [4]: Schunck's constraint-line clustering algorithm [19] uses cluster analysis to determine the dominant motion of a given image region. Black and Anandan [7] use robust estimators for the same purpose. However, approaches focusing on dominant motion do not explicitly form a multiple motion model as they only provide single estimates where many motions may prevail. Alternatively, a number of authors have studied inhibitory smoothness constraints [2, 15, 16, 17] which relax smoothness requirements at image regions exhibiting high spatial 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 [18]. Again, such schemes are limited by an explicit single motion model.

Schemes for estimating optical flow at regions of partial translucency have also appeared: Bergen *et al.* [6] use a multiple motion model which iteratively refines sequences of difference images, each one containing a single motion. In the case of spatiotemporal frequency models, such as Fleet's [9], translucency could be handled with an appropriate constraint-integration algorithm. Schemes capable of handling both occlusion and translucency have also been published: Irani *et al.* [12] use a sequential tracking and registration algorithm to independently compute the velocity of possibly translucent objects undergoing different motions. Multilayer and superposition models also provide paradigms for the estimation of multiple motions. Shizawa and Mase [20] propose a computational framework to perform multiple motion analysis by formulating a superposition model from which constraints inherent to multiple motions can be derived. Wang and Adelson [22] use a clustering algorithm to group velocities into layers, each consistent with an affine model of motion. Darrell and Pentland's algorithm [8] perform analysis of multiple motions by comput-

ing a minimal set of layers, each describing a coherent motion from the scene, which yield a complete description of motion when superposed. Jepson and Black [13] propose mixture models to segment multiple motions with the use of an EM algorithm.

While these approaches constitute important contributions, they often lack the ability to determine the nature of the image events giving rise to multiple motions: Occlusion may not be differentiated from translucency and it is difficult to identify the motions of both the occluding and occluded surfaces. In this paper, we demonstrate, under a set of reasonable assumptions, that such distinctions can be made through a Fourier analysis of these image events and 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 3-d translation $\mathbf{T} = (T_x, T_y, T_z)$ and rotation $\mathbf{\Omega} = (\Omega_x, \Omega_y, \Omega_z)$:

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

where $\mathbf{x}^T = (x, y)$ is the perspective projection of a point $\mathbf{P} = (X, Y, Z)$ 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) when 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 studying these phenomena in the frequency domain arises from a number of observations:

- The structure of the frequency distortion generated by occlusion in the frequency domain is unknown [9] and needs to be investigated[1];
- Frequency-based optical flow methods possess the advantage of providing several constraints for image regions of interest, thus allowing the detection of multiple motions, given adequate constraint-grouping methods [9, 13];
- High accuracy in the estimation of image motion may be obtained with phase-based filtering methods operating in the frequency domain [3];
- The study of phenomena such as occlusion and translucency is simplified in the frequency domain.

In addition, the hypotheses posed towards studying occlusion and translucency are that velocities are locally constant and individual signals satisfy Dirichlet conditions.

2.1 Occlusion in the Frequency Domain

The Fourier transform of the optical flow constraint equation is

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

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 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:

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

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 [9]. 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[2].

Following Fleet [9], 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

[1] Non-Fourier motions have been studied by Fleet and Langley [10].

[2] The aperture problem arises when the Fourier spectrum of $\mathbf{I}(\mathbf{x})$ is concentrated on a line rather than on a plane [9, 14]. 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.

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:

$$\begin{aligned} \mathbf{I}(\mathbf{x},t) &= \mathbf{I}_1(\mathbf{x}-\mathbf{v}_1 t) \\ &+ [1-\mathbf{U}(\mathbf{x}-\mathbf{v}_1 t)]\,\mathbf{I}_2(\mathbf{x}-\mathbf{v}_2 t). \end{aligned} \quad (4)$$

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

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

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 the image events giving rise to multiple motions is shown.

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: $\mathbf{I}(x,t)=\mathbf{I}_1(x-v_1t)$. 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:

$$\begin{aligned} \mathbf{I}(x,t) &= \mathbf{u}(x-v_1t)\mathbf{I}_1(x-v_1t) \\ &+ (1-\mathbf{u}(x-v_1t))\mathbf{I}_2(x-v_2t), \end{aligned} \quad (6)$$

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:

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

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 $k_1=2\pi f_1$ and $k_2=2\pi f_2$ and let $\mathbf{I}_1(x-v_1t)=a_1\cos(k_1x-v_1t)$ and $\mathbf{I}_2(x-v_2t)=a_2\cos(k_2x-v_2t)$. Then the frequency spectrum of the occlusion obtained by substituting $\mathbf{I}_1(x)$ and $\mathbf{I}_2(x)$ into (6) is:*

$$\begin{aligned} \hat{\mathbf{I}}(k,\omega) &= \frac{\pi a_1}{2}\delta(k\pm k_1,\omega\mp k_1v_1) \\ &+ \frac{(1-\pi)a_2}{2}\delta(k\pm k_2,\omega\mp k_2v_2) \\ &+ \frac{i}{2}\left(\frac{a_2\delta(kv_1+\omega\pm k_2\Delta v)}{(k\pm k_2)}\right. \\ &- \left.\frac{a_1\delta(kv_1+\omega)}{(k\pm k_1)}\right). \end{aligned} \quad (8)$$

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 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 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 [11].

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:*

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

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

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

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. However, the distortion terms are entirely 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:*

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

where $\mathbf{n} = (n_x, n_y)$ *are integers,* $\mathbf{x} = (x, y)$ *are spatial coordinates,* $\mathbf{k}_1 = (n_x k_{x1}, n_y k_{y1})$ *and* $\mathbf{k}_2 = (n_x k_{x2}, n_y k_{y2})$ *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:*

$$\begin{aligned} &\hat{\mathbf{I}}(\mathbf{k},\omega) = \\ &\quad \pi \sum_{\mathbf{n}=-(\infty,\infty)}^{(\infty,\infty)} c_{1\mathbf{n}}\delta(\mathbf{k} - \mathbf{k}_1, \omega + \mathbf{k}_1^T \mathbf{v}_1) \\ &+ (1-\pi) \sum_{\mathbf{n}=-(\infty,\infty)}^{(\infty,\infty)} c_{2\mathbf{n}}\delta(\mathbf{k} - \mathbf{k}_2, \omega + \mathbf{k}_2^T \mathbf{v}_2) \\ &+ i \sum_{\mathbf{n}=-(\infty,\infty)}^{(\infty,\infty)} \left(\frac{c_{2\mathbf{n}}\delta(\mathbf{k}^T\mathbf{v}_1 + \omega - \mathbf{k}_2^T\Delta\mathbf{v})}{(\mathbf{k} - \mathbf{k}_2)^T\vec{\eta}} \right. \\ &- \left. \frac{c_{1\mathbf{n}}\delta(\mathbf{k}^T\mathbf{v}_1 + \omega)}{(\mathbf{k} - \mathbf{k}_1)^T\vec{\eta}} \right), \end{aligned} \tag{13}$$

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 and the distortion terms are contained in the imaginary part of the Fourier spectrum[3].

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 [9]. Under the local assumption of spatially constant $f(\rho_1)$, with translucency

[3] For details concerning the proofs of these propositions, the interested reader is referred to [5].

factor φ, (14) is reformulated as a weighted superposition of intensity profiles, written as

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

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:

$$\begin{aligned}\hat{\mathbf{I}}(\mathbf{k},\omega) = \\ \varphi \sum_{\mathbf{n}=-(\infty,\infty)}^{(\infty,\infty)} c_{1\mathbf{n}}\delta(\mathbf{k}-\mathbf{k}_1, \omega+\mathbf{k}_1^T\mathbf{v}_1) \\ + \quad (1-\varphi)\sum_{\mathbf{n}=-(\infty,\infty)}^{(\infty,\infty)} c_{2\mathbf{n}}\delta(\mathbf{k}-\mathbf{k}_2, \omega+\mathbf{k}_2^T\mathbf{v}_2).\end{aligned} \quad (16)$$

With the exception of the imaginary 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 imaginary distortion terms vanish.

3.5 Geometric Interpretation

In the simplest case involving sinusoidal signals, (8) shows that the frequency spectra of the profiles $\mathbf{I}_1(x)$ and $\mathbf{I}_2(x)$ are preserved to within scaling factors under occlusion. In addition, the imaginary term of (8) represents the distortion created by the occluding boundary. Figure 1a shows a Gaussian-smoothed, continuous frequency spectrum of (8) for an occluding 1-d cosine function with spatial frequency $k_1 = \frac{2\pi}{16}$ and velocity $v_1 = 1$. The occluded cosine function has spatial frequency $k_2 = \frac{2\pi}{8}$ and velocity $v_2 = -1$. The horizontal axis represents spatial frequency while the vertical axis is temporal frequency. The peaks located at $\pm(k_1, -k_1 v_1)$ and $\pm(k_2, -k_2 v_2)$ depict the spatiotemporal frequencies of the signals and fit the constraint lines $kv_1 + \omega = 0$ and $kv_2 + \omega = 0$. The oblique spectra intersecting the peaks represent the distortion due to the occlusion boundary and fit the constraint lines $kv_1 + \omega \pm k_2 v_2 = 0$ and $kv_1 + \omega = 0$. Geometrically, these lines are parallel to the constraint line of the occluding signal.

Figure 1b shows the frequency spectrum for the case of an occluded signal exhibiting many spatial frequencies. The structure of the spectrum is a generalization of the simplest case: For every frequency nk_1 and nk_2, an oblique spectrum originating from the occlusion boundary intersects $(k_2, -nk_2 v_2)$ and $(k_1 - nk_1 v_1)$. These spectra fit the constraint lines $kv_1 + \omega - nk_2\Delta v = 0$ and $kv_1 + \omega = 0$ associated with both signals. The occluding 1-d sinusoidal has spatial frequency $k_1 = \frac{2\pi}{16}$ and velocity $v_1 = 1$, while the occluded signal has spatial frequencies $nk_2 = 2\pi(\frac{n}{8}), n = 1,2,3$ and velocity $v_2 = -1$. This continuous spectrum was generated with a Gaussian-smoothed version of (10).

The 2-d generalization (13) of (10) is also similar in terms of its geometric interpretation: The constraint lines $kv_1 + \omega = 0$, $kv_2 + \omega = 0$ and $kv_1 + \omega - nk_2\Delta v = 0$ from (10), in which the Fourier spectra of both the occluding and occluded signals fit, become constraint planes in (13). The sets of discrete frequency locations $\{(\mathbf{k}_1, -\mathbf{k}_1^T\mathbf{v}_1)\}$ and $\{(\mathbf{k}_2, -\mathbf{k}_2^T\mathbf{v}_2)\}$ respectively fit the planes $\mathbf{k}^T\mathbf{v}_1 + \omega = 0$ and $\mathbf{k}^T\mathbf{v}_2 + \omega = 0$. In the imaginary distortion term, the Dirac δ functions $\delta(\mathbf{k}^T\mathbf{v}_1 + \omega - \mathbf{k}_2^T\Delta\mathbf{v})$ and $\delta(\mathbf{k}^T\mathbf{v}_1 + \omega)$ 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 $\mathbf{k}_1$ and $\mathbf{k}_2$ exhibited by both signals, a frequency spectrum fitting the planes given by $\mathbf{k}^T\mathbf{v}_1 + \omega - \mathbf{k}_2^T\Delta\mathbf{v} = 0$ and $\mathbf{k}^T\mathbf{v}_1 + \omega = 0$ is found. The magnitudes of these planar Fourier spectra are determined by their corresponding scaling functions $c_{2\mathbf{n}}[(\mathbf{k}-\mathbf{k}_2)^T\vec{\eta}\ ]^{-1}$ and $c_{1\mathbf{n}}[(\mathbf{k}-\mathbf{k}_1)^T\vec{\eta}\ ]^{-1}$. Hence, (13) provides a number of useful constraint planes: 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

Several experiments were performed in order to verify the accuracy of the propositions. The Fourier spectra obtained with both a standard FFT algorithm and those predicted by the theory were compared. In addition, the effects of various levels of zero-mean Gaussian noise on these spectra were visually examined.

4.1 Verification

In order to verify the propositions against numerical experiments, two 1-d sinusoidals which respectively act as occluding and occluded surfaces were used. Expression (6) is used with $\mathbf{I}_1(x - v_1 t) = a_1\cos(k_1 x - v_1 t)$ and $\mathbf{I}_2(x - v_2 t) = a_2\cos(k_2 x - v_2 t)$, where $\mathbf{I}_1$ and $\mathbf{I}_2$ are the occluding and occluded surfaces with respective frequencies $k_1 = \frac{2\pi}{16}$ and

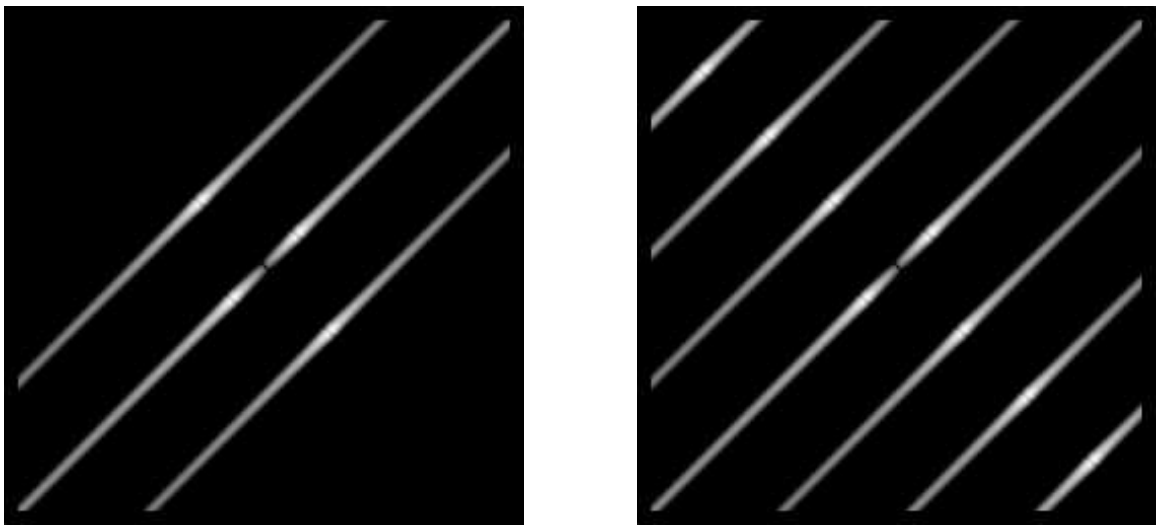

Figure 1: **a) (left):** *Gaussian-smoothed frequency spectrum produced with (8). The occluding signal has frequency* $k_1 = \frac{2\pi}{16}$ *and velocity* $v_1 = 1$*, while the occluded signal has frequency* $k_2 = \frac{2\pi}{8}$ *and velocity* $v_2 = -1$*.* **b) (right):** *Gaussian=smoothed frequency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

$k_2 = \frac{2\pi}{8}$ and velocities $v_1 = 1$ and $v_2 = -1$. Constants a_1 and a_2 correspond to signal amplitudes. To limit boundary conditions when numerically computing fast Fourier transforms, the image was windowed with a Gaussian envelope. The result, shown in Figure 2a, is analytically expressed as

$$\mathbf{I}(x,t)G(x,t;\sigma), \tag{17}$$

where

$$\begin{aligned} \mathbf{I}(x,t) &= \mathbf{u}(x - v_1 t)a_1 \cos(k_1 x - v_1 t) \\ &+ (1 - \mathbf{u}(x - v_1 t))a_2 \cos(k_2 x - v_2 t) \end{aligned}$$

and

$$G(x,t;\sigma) = \frac{1}{2\pi\sigma^2} e^{-\left(\frac{x^2+y^2}{2\sigma^2}\right)}.$$

A standard deviation of 40 is used for the windowing of the signal. The discrete Fourier transform of (17) 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. Analytically, the continuous Fourier transform of (17) is

$$\hat{\mathbf{I}}(k,\omega) * \hat{G}(k,\omega;\sigma), \tag{18}$$

where $\hat{\mathbf{I}}(k,\omega)$ is proposition 1 and $\hat{G}(k,\omega;\sigma) = e^{-\frac{1}{2}\sigma^2(k^2+\omega^2)}$. Expression (18) 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[4].

4.2 Gaussian Noise

It is important to determine the extent of disruption noise may cause to the frequency structure of occlusion phenomena. Towards this end, the signal of Figure 2a was corrupted with various levels of zero-mean Gaussian noise exhibiting standard deviations of 12, 25 and 40[5]. Figures 3a, b, and c illustrate the resulting fast Fourier transforms, obtained with a standard FFT algorithm for each one of these noise levels. Although the frequency structure remains visible for all noise levels, one may observe a progressive deterioration of the original signal.

[4] Note the aliasing effect in the upper-left and lower-right corner of Figure 2b which are not found in Figure 2c.

[5] These standard deviations are expressed with respect to grayvalues. They approximately correspond to 5, 10 and 15 percent of signal amplitude.

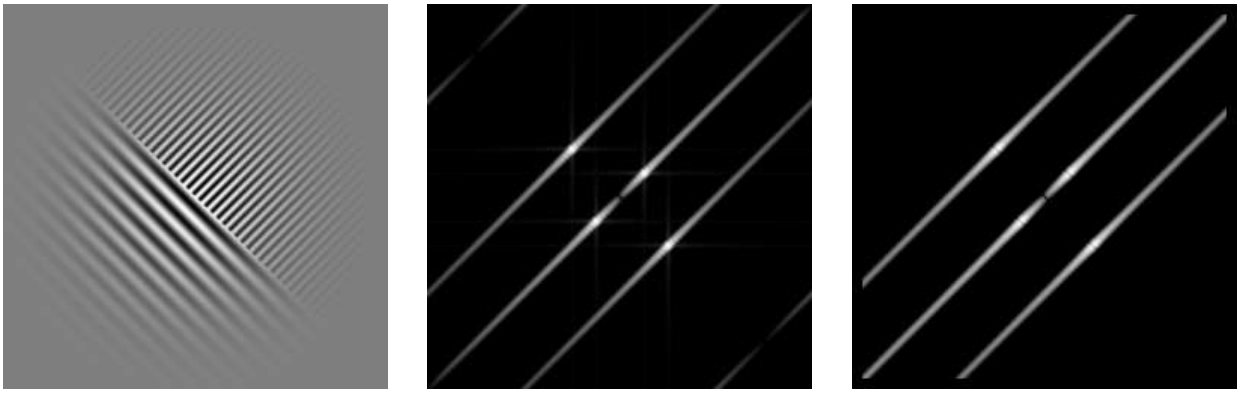

Figure 2: **a) (left):** *Gaussian-windowed signal with sinusoidals acting as occluding and occluded surfaces. Occluding signal: Frequency* $k_1 = \frac{2\pi}{16}$ *and velocity* $v_1 = 1$. *Occluded signal: Frequency* $k_2 = \frac{2\pi}{8}$ *and velocity* $v_2 = -1$. **b) (middle):** *Fourier spectrum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

5 Summary

Under a set of reasonable hypotheses, such as locally constant velocity and intensity profiles satisfying Dirichlet conditions, the structure of occlusion in the frequency domain was determined and its relation to translucency was shown. Several advantages are brought by the frequency analysis of occlusion:

- The knowledge of the frequency structure of the distortion arising from occlusion allows an eventual detection mechanism to locally determine the causes of multiple motions: The presence of the distortion indicates an occlusion while its absence signifies translucency;
- The geometric parallelism between the structure of the distortion and the frequency spectrum of the occluding intensity profile allow us to locally differentiate the occluding surface from the occluded, without the need of any knowledge concerning the structure of the scene;
- This analysis forms a basis for an adequate constraint-grouping algorithm capable of distinguishing occlusion from translucency and identifying occluding and occluded surfaces. In addition, a further reduction in signal-to-noise ratios is likely to result from including distortion in models of multiple motions.

Future research in this area includes the construction of models of multiple motions which account for the structure of image events such as occlusion and translucency. In particular, one may reformulate multilayer and mixture models [9, 13], in order to obtain crucial information concerning image events leading to multiple motions.

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] J. Aisbett. Optical flow with intensity-weighted smoothing. *IEEE PAMI*, 11(5):512–522, 1989.

[3] J. L. Barron, D. J. Fleet, and S. S. Beauchemin. Performance of optical flow techniques. *IJCV*, 12(1):43–77, 1994.

[4] S. S. Beauchemin and J. L. Barron. The computation of optical flow. Technical Report TR-450, Dept. of Computer Science, Univ. of Western Ontario, March 1995.

[5] S. S. Beauchemin and J. L. Barron. The structure of occlusion in fourier space. In *Vision Interface*, pages 112–119, Quebec City, Canada, May 1995.

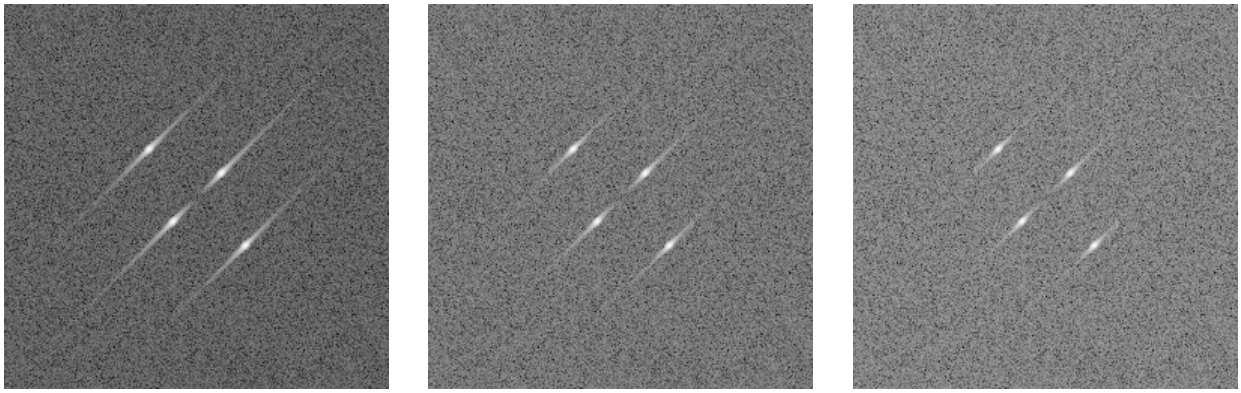

Figure 3: *from left to right: Fourier spectra of signal corrupted with zero-mean Gaussian noise exhibiting standard deviations of 12, 25 and 40, obtained with a standard FFT algorithm.*

[6] 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.

[7] 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.

[8] 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.

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

[10] D. J. Fleet and K. Langley. Computational analysis of non-fourier motion. *Vision Research*, 34(22):3057–3079, 1995.

[11] J. D. Gaskill. *Linear Systems, Fourier Transforms and Optics*. Wiley & Sons, Inc., 1978.

[12] M. Irani, B. Rousso, and S. Peleg. Recovery of egomotion using image stabilization. In *CVPR*, pages 454–460, Seattle, Washington, June 1994.

[13] 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.

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

[15] H.-H. Nagel. Constraints for the estimation of vector fields from image sequences. In *Proceedings of IJCAI*, pages 945–951, Karlsruhe, Federal Republic of Germany, August 1983.

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

[17] H.-H. Nagel. On the estimation of optical flow: Relations between different approaches and some new results. *Artificial Intelligence*, 33:299–324, 1987.

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

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

[20] 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.

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

[22] J. Y. A. Wang and E. H. Adelson. Layered representation for motion analysis. In *Proceedings of CVPR*, 1993.