Decoupling Fourier Frequency Components in Dynamic Image Sequences containing Multiple Moving Objects

David Vernon
Department of Computer Science
Maynooth College
Ireland

Abstract

One of the principal problems associated with the use of Fourier techniques in the analysis of image sequences comprising moving objects is that different objects, including the background which may be viewed as a complex object moving with zero velocity, may share a common spectral support (*i.e.* comprise identical spatial frequencies). The Fourier domain represents these multiple frequency components by a single frequency (k_x, k_y) with phase equal to the sum of the phases of the individual components and with a magnitude which is dependent on the relative phase of these components. In this paper, we show how the individual frequency components which are associated with each moving object (or, rather, each set of objects moving with distinct constant velocity) can be computed from the single compound frequency magnitude and phase.

Introduction

The Fourier transform can be used to compute the velocity of moving images by exploiting the shift property of the transform whereby an image which undergoes a fixed displacement exhibits only a frequency-dependent phase change in the Fourier domain [1, 2, 3, 4, 5, 6]. This technique has been extended to allow the computation of the velocity of a distinct moving object in image sequences [7]. Recently, the feasibility of segmenting moving objects in image sequences by isolating coherent wave profiles in images has been demonstrated [8]. In this case, image sequences which are temporally-changing 2-D signals are treated as moving - and changing - 2D wave profiles. Each object in the image can also be viewed as a distinct moving wave profile. As noted above, the Fourier components of wave profiles which move with constant velocity exhibit a regular frequency-dependent phase change. A phase-velocity Hough transform [9] isolates coherent wave profiles by grouping Fourier components exhibiting a phase change which is consistent with a given velocity and segmentation is effected by re-constituting the moving wave profile - the object - from the grouped Fourier components by filtering out all other components and then computing the inverse Fourier transform. In this paper, we are concerned with highlighting the problems which attend the use of the Fourier transform and we show how these problems may be overcome.

Background

The approach set out in [7] and [8] depends on a wave-based model of the dynamic image. Specifically, an image $g(x, y, t)$ is regarded as a time-varying surface and an object O_i which is moving in the image is viewed as a wave with characteristic shape propagating through the image space with a velocity $v_i(t)$. That is, an individual wave profile arises from the movement of a single object O_i, translating with constant velocity (or, rather, a group of objects moving with uniform constant velocity). This group could, of course, be the background with, *e.g.*, zero velocity. Such a wave profile satisfies the 2-D differential wave equation. Thus:

$$\frac{\partial^2 \psi^i(x,y,t)}{\partial x^2} + \frac{\partial^2 \psi^i(x,y,t)}{\partial y^2} = \frac{1}{v_i} \frac{\partial^2 \psi^i(x,y,t)}{\partial t^2}$$

The solution of this wave equation $\psi^i(x,y,t)$ is a description of the object as a grey-level wave profile propagating with constant velocity, *i.e.*,

$$\psi^i(x,y,t) = f^i(x - v_{x_i}t, y - v_{y_i}t).$$

By the principle of linear superposition, the wave corresponding to a given object, $\psi^i(x,y,t) = f^i(x - v_{x_i}t, y - v_{y_i}t)$, can be decomposed into constituent components:

$$\psi^i(x,y,t) = \sum_{k=1}^{l_i} c_k^i f_k^i(x - v_{x_k}^i t, y - v_{y_k}^i t)$$

Similarly, the total optical field can be decomposed:

$$\psi(x,y,t) = \sum_{j=1}^{n} c_j f_j\,(x - v_{x_j} t, y - v_{y_j} t)$$

In addition, the total optical field is the sum of all *m* of the individual object waves:

$$\psi(x,y,t) = \sum_{i=1}^{m} \sum_{k=1}^{l_i} c_k^i f_k^i(x - v_{x_k}^i t, y - v_{y_k}^i t)$$

The task, then, is to decompose the image into constituent components $c_j f_j\,(x - v_{x_j} t, y - v_{y_j} t)$ and to group them into *m* sets $c_k^i f_k^i(x - v_{x_k}^i t, y - v_{y_k}^i t)$, $1 \le i \le m$, such that $(v_{x_k}^i, v_{y_k}^i)$ is constant. As noted above, the Fourier transform is used to accomplish the decomposition and the a phase-velocity Hough transform is used to effect the grouping [7, 8]. This decomposition and grouping allows one to identify uniquely the velocity of each moving wave profile, *i.e.*, each object or group of objects and its constituent components which can then be used to reconstitute that object. We observe here, however, that a given component $f_j\,(x,y)$ may well be present in more than one object wave, although it will have a different velocity $(v_{x_k}^i, v_{y_k}^i)$ and a different coefficient c_k^i for each object wave ψ^i.

Implications of using the Fourier Transform

In the discrete case, the inverse Fourier transform is given by:

$$\begin{aligned}\psi(x,y) &= \frac{1}{(2\pi)^2} \sum_{k_x} \sum_{k_y} \left|\mathrm{F}(k_x,k_y)\right| e^{i\phi(k_x,k_y)} e^{-i(k_x x + k_y y)} \\ &= \mathcal{F}^{-1}(\mathrm{F}(k_x,k_y))\end{aligned}$$

That is, $\psi(x,y)$ can be constructed from a linear combination of elementary functions having the form $e^{-i(k_x x + k_y y)}$, each appropriately weighted in amplitude and phase by a complex factor $\mathrm{F}(k_x,k_y)$. Thus, in the wave model being adopted, $c_j = \mathrm{F}(k_x,k_y)$ and $f_j(x,y) = e^{-i(k_x x + k_y y)}$. That is, the constituent components are exponential functions of the form $e^{-i(k_x x + k_y y)}$ with complex coefficients $\mathrm{F}(k_x,k_y) = \mathcal{F}(\psi(x,y))$. We observed in the previous section that several object wave profiles ψ^i may comprise the same components $f_j(x,y) = e^{-i(k_x x + k_y y)}$ albeit with different coefficients c_k^i. That is, the model assumes that each coefficient c_k^i is explicitly represented. In the case of the exponential components of the Fourier domain, this is equivalent to an assumption that, for a given component, there is a distinct complex coefficient $\mathrm{F}^i(k_x,k_y)$ for each constituent object wave profile. Unfortunately, this is not the case: there is only one unique coefficient $\mathrm{F}(k_x,k_y) = \mathcal{F}(\psi(x,y))$ for each component $e^{-i(k_x x + k_y y)}$. This limitation of the Fourier transform causes significant problems for any Fourier-based object velocity estimation and segmentation procedure since such an approach requires the distinct coefficients $c_k^i = \mathrm{F}^i(k_x,k_y)$ for each wave profile ψ^i rather than the conflated coefficient $c_j = \mathrm{F}(k_x,k_y)$. The goal of this paper is to investigate this limitation, to see whether or not it is possible to determine each $\mathrm{F}^i(k_x,k_y)$ (equivalently c_k^i) from $\mathrm{F}(k_x,k_y)$ (equivalently c_j) and, if it is possible, to identify the circumstances under which it is possible.

$$\mathsf{F}_2(k_{x_i},k_{y_i}) \quad = \quad \mathsf{F}^B(k_{x_i},k_{y_i}) + Rot(\mathsf{F}^{O_1}(k_{x_i},k_{y_i}),\delta\phi_1) \qquad (2)$$

Consequently, if we know the velocity (v_x, v_y) and hence the phase shift $\delta\phi_1$, and knowing $\mathsf{F}_1(k_{x_i},k_{y_i})$ and $\mathsf{F}_2(k_{x_i},k_{y_i})$, we can solve equations (1) and (2) simultaneously for $\mathsf{F}^B(k_{x_i},k_{y_i})$ and, in particular, for $\mathsf{F}^{O_1}(k_{x_i},k_{y_i})$. Subtracting (1) from (2) we have:

$$\mathsf{F}_2(k_{x_i},k_{y_i}) - \mathsf{F}_1(k_{x_i},k_{y_i}) \quad = \quad Rot(\mathsf{F}^{O_1}(k_{x_i},k_{y_i}),\delta\phi_1) - \mathsf{F}^{O_1}(k_{x_i},k_{y_i}) \qquad (3)$$

We can solve this analytically using, *e.g.*, homogeneous coordinates to represent the vectors and the rotation. In the case where there are two objects moving, in addition to a stationary object or background, and assuming that the velocity of each object is known or can be estimated, we can recover the components of each individual object in a similar manner but this time exploiting three images rather than two.

Thus, we see that it is indeed possible to determine each $\mathsf{F}^i(k_x,k_y)$ corresponding to a given moving object or wave profile from $\mathsf{F}(k_x,k_y)$, as required, if the velocity of each object or wave profile is known. This is the crucial point: the identification of the object velocity. We noted above that it is possible to identify the velocity of a distinct moving object in an image sequence [7]. However, this paper did not address the computation of velocity in images where there are several objects moving with different velocities. In the future, we will show how it can be extended to images containing two objects (and, in principle at least, several objects).

References

[1] Jacobson and H. Wechsler, 'Derivation of optical flowusing a spatiotemporal-frequency approach', *Computer Vision, Graphics, and Image Processing*, **38**, 29-65 (1987).

[2] M.P. Cagigal, L. Vega, P. Prieto, 'Object movement characterization from low-light-level images', *Optical Engineering*, **33(8)**, 2810-2812 (1994).

[3] M.P. Cagigal, L. Vega, P. Prieto, 'Movement characterization with the spatiotemporal Fourier transform of low-light-level images", *Applied Optics*, **34(11)**, 1769-1774 (1995).

[4] S. A. Mahmoud, M.S. Afifi, and R. J. Green, 'Recognition and velocity computation of large moving objects in images", *IEEE Transactions on Acoustics, Speech, and Signal Processing*, **36(11)**, 1790-1791 (1988).

[5] S. A. Mahmoud, 'A new technique for velocity estimation of large moving objects', *IEEE Transactions on Signal Processing*, **39(3)**, 741-743 (1991).

[6] S.A. Rajala, A. N. Riddle, and W.E. Snyder, 'Application of one-dimensional Fourier transform for tracking moving objects in noisy environments", *Computer Vision, Graphics, and Image Processing*, **21**, 280-293 (1983).

[7] D. Vernon, 'Phase-Based Measurement of Object Velocity in Image Sequences using the Hough Transform', *Optical Engineering*, (in press).

[8] D. Vernon, 'Segmentation in Dynamic Image Sequences by isolation of coherent wave profiles', Proceedings of the 4th European Conference on Computer Vision, Springer-Verlag, 293-303 (1996)

[9] P.V.C. Hough, 'Method and Means for Recognising Complex Patterns' U.S. Patent 3,069,654, (1962).

Development

The first point to note is that $c_j = \sum_{i=1}^{m} c_k^i$. This can be demonstrated in a straightforward manner as follows.

We have $\psi = \sum_{i=1}^{m} \psi^i$ and, thus, $\mathcal{F}(\psi) = \mathcal{F}(\sum_{i=1}^{m} \psi^i)$. Since the Fourier transform is a linear operation, we have $\mathcal{F}(\psi) = \mathcal{F}(\sum_{i=1}^{m} \psi^i) = \sum_{i=1}^{m} \mathcal{F}(\psi^i)$. Equivalently, $\mathrm{F}(k_x, k_y) = \sum_{i=1}^{m} \mathrm{F}^i(k_x, k_y)$ or $c_j = \sum_{i=1}^{m} c_k^i$, for any given Fourier component k. Essentially, the Fourier component $\mathrm{F}(k_x, k_y)$ is the vector sum of each $\mathrm{F}^i(k_x, k_y)$ and its resultant magnitude and phase will be dependent on the magnitudes and phases of the individual components $\mathrm{F}^i(k_x, k_y)$. Thus, for a given component (equivalently, a given spatial frequency (k_x, k_y) and a given exponential component defined by $e^{-i(k_x x + k_y y)}$), the Fourier transform only gives us $c_j = \mathrm{F}(k_x, k_y)$, the composite coefficient, rather than what we want, *i.e.*, the coefficient c_k^i corresponding to each wave profile ψ^i . Our goal in this paper is to demonstrate how each $\mathrm{F}^i(k_x, k_y)$ can be determined from $\mathrm{F}(k_x, k_y)$ under certain circumstances.

Consider a waveform $\psi(x - v_x \delta t, y - v_y \delta t)$, which is $\psi(x, y)$ shifted by $(v_x \delta t, v_y \delta t)$, *i.e.*, it is travelling with constant velocity (v_x, v_y) . The shift property of the Fourier transform states that the Fourier transform of any such shifted function is given by:

$$
\begin{aligned}
\mathcal{F}(\psi(x - v_x \delta t, y - v_y \delta t)) &= \left|\mathrm{F}(k_x, k_y)\right| e^{i\phi(k_x, k_y)} e^{-i(k_x v_x \delta t + k_y v_y \delta t)} \\
&= \mathcal{F}(\psi(x, y)) e^{-i(k_x v_x \delta t + k_y v_y \delta t)}
\end{aligned}
$$

That is, and as we have already noted, a spatial shift of a signal only produces a frequency-dependent phase change in the Fourier domain. The consequence of this is that the components of ψ^i translating with uniform velocity only undergo a phase change, with no change in magnitude; that is, the vector $\mathrm{F}^i(k_x, k_y)$ is only rotated by an angle given by the phase change $-(k_x v_x \delta t + k_y v_y \delta t)$.

Consider now the situation where we have an image of a scene comprising one moving object and a stationary background. Let ψ_{t_1} be this image at time t_1 and let ψ_{t_2} be this image at time t_2 . Let $\mathrm{F}_1(k_x, k_y) = \mathcal{F}(\psi_{t_1}(x, y))$ and let $\mathrm{F}_2(k_x, k_y) = \mathcal{F}(\psi_{t_2}(x, y))$. For any given spatial frequency (k_{x_i}, k_{y_i}) , we have a frequency components $\mathrm{F}_1(k_{x_i}, k_{y_i})$ and $\mathrm{F}_2(k_{x_i}, k_{y_i})$ at times t_1 and t_2 , respectively. Each frequency component has a component due to the background $\mathrm{F}^B(k_{x_i}, k_{y_i})$ and the object moving in the image $\mathrm{F}^{O_1}(k_{x_i}, k_{y_i})$ (assuming that they share a common frequency support and that (k_{x_i}, k_{y_i}) is a member of this support). Thus,

$$
\mathrm{F}_1(k_{x_i}, k_{y_i}) = \mathrm{F}^B(k_{x_i}, k_{y_i}) + \mathrm{F}^{O_1}(k_{x_i}, k_{y_i}) \tag{1}
$$

Now, if the background is stationary and the object is translating with constant velocity then $\mathrm{F}^B(k_{x_i}, k_{y_i})$ is constant and $\mathrm{F}^{O_1}(k_{x_i}, k_{y_i})$ is subject to a phase shift $\delta\phi_1 = -(k_{x_i} v_x \delta t + k_{y_i} v_y \delta t)$. Thus,