Chapter 6

Image Velocity as Local Phase Behaviour

This chapter is intended to motivate and introduce a phase-based definition of image velocity, and a corresponding technique for its measurement. To begin, let $R(x, t)$ denote the response of a velocity-tuned band-pass filter. Because the filter kernels, such as $Gabor(x, t; k_0, \omega_0, C)$ are complex-valued, $R(x, t)$ is also complex-valued and can therefore be expressed as

$$ R(x, t) = \rho(x, t) e^{i\phi(x, t)}, \quad (6.1) $$

where $\rho(x, t)$ and $\phi(x, t)$ denote its amplitude and phase components (1.4):

$$ \rho(x, t) = |R(x, t)|, $$
$$ \phi(x, t) = \text{arg}[R(x, t)]. $$

In the search for an appropriate definition of image velocity, our goal is to determine which properties of $R(x, t)$ evolve in time according to the motion field. The thesis set forth in this monograph purports that the phase response $\phi(x, t)$ is such a property. In particular, it is shown that the temporal evolution of (spatial) contours of constant phase provides a better approximation to the motion field than do contours of constant amplitude $\rho(x, t)$ or the level contours of $R(x, t)$.

We begin with a demonstration of the robustness of $\phi(x, t)$ as compared with $\rho(x, t)$ in Section 6.1. Section 6.2 then introduces the phase-based definition of component velocity. Section 6.3 shows the connection between this definition and local frequency analysis, and Section 6.4 discusses its relationship to the definitions implicit in other techniques.

D. J. Fleet, Measurement of Image Velocity
6.1 Why Phase Information?

If the temporal variation of image intensity was due solely to image translation, as in \( I(x, t) = I(x - vt, 0) \), then it is easy to show that the filter outputs would also translate, as in \( R(x, t) = R(x - vt, 0) \). Moreover, we would therefore expect that, in principle, all the techniques described in Chapter 5 should produce accurate estimate of the velocity \( v \), so long as the aperture problem and undersampling of the signal are not too severe.

However, image translation is only a rough approximation to the typical time-varying behaviour of image intensity. As discussed in Chapter 2, a more realistic model includes contrast variation and geometric deformation due to perspective projection; and it is from this standpoint that we propose the use of phase information. In particular, we argue that the evolution of phase contours provides a better approximation to the projected motion field than the filter response \( R(x, t) \), in that the amplitude of response \( \rho(x, t) \) is generally very sensitive to changes in contrast and to local variations in the scale, speed, and orientation of the input.

We demonstrate the robustness of phase compared to amplitude using several 1-d examples which serve to approximate the dilation of an image as a camera approaches a planar surface. In particular, we consider the time-varying image

\[
I(x, t) = I(x(1 - \alpha t), 0),
\]

for some \( \alpha > 0 \). The initial pattern \( I_0(x) \equiv I(x, 0) \) is simply stretched as \( t \) increases. The velocity field for this deformation is given by the motion of fixed points, say \( \xi \), in the 1-d pattern \( I_0(x) \). In image coordinates these points appear on paths generated by \( x(1 - \alpha t) = \xi \). These paths are clearly visible from the inputs in Figures 6.1 and 6.2 (top-left).

Figure 6.1 (top) shows the time-varying intensity pattern generated by (6.2) for \( I_0(x) = \sin(2\pi f_0 x) \), and the time-varying response of the real part of a Gabor filter (4.4) tuned to spatial frequency \( 2\pi f_0 \) and to zero velocity (vertically oriented structure in Figure 6.1). The amplitude and phase components of \( R(x, t) \) are shown, as functions of space and time, in Figure 6.1 (middle). Finally, the bottom panels show the level contours of constant amplitude and constant phase superimposed upon the input. While the phase contours provide a good approximation to the motion field, the amplitude contours do not. The reasons for this amplitude behaviour are straightforward: The behaviour of \( \rho(x, t) \) reflects how well the local input structure matches the filter tuning. In this case, with the amplitude of the input sinusoid held constant, they match very well in the centre of the image and less so as one moves out from the centre. Other things being equal, \( \rho(x, t) \) increases for inputs closer to the principal frequency to which the filter is