Chapter 5

Local Constraints for the Perception of Binocular 3D Motion

Martin Lages
University of Glasgow, UK

Suzanne Heron
University of Glasgow, UK

Hongfang Wang
University of Glasgow, UK

ABSTRACT

The authors discuss local constraints for the perception of three-dimensional (3D) binocular motion in a geometric-probabilistic framework. It is shown that Bayesian models of binocular 3D motion can explain perceptual bias under uncertainty and predict perceived velocity under ambiguity. The models exploit biologically plausible constraints of local motion and disparity processing in a binocular viewing geometry. Results from computer simulations and psychophysical experiments support the idea that local constraints of motion and disparity processing are combined late in the visual processing hierarchy to establish perceived 3D motion direction.

1. INTRODUCTION

The perceptual inference of the three-dimensional (3D) external world from two-dimensional (2D) retinal input is a fundamental problem (Berkeley, 1709/1975; von Helmholtz, 1910/1962) that the visual system has to solve through neural computation (Poggio, Torre, & Koch, 1985; Pizlo, 2001). This is true for static scenes as well as for dynamic events. For dynamic events the inverse problem implies that the visual system estimates motion in 3D space from local encoding and spatio-temporal processing.

Velocity in 3D space is described by motion direction and speed. Motion direction can be measured in terms of azimuth and elevation angle, and motion direction together with speed is conveniently expressed as a vector in a 3D Cartesian
Local Constraints for the Perception of Binocular 3D Motion

coordinate system. Estimating local motion vectors is highly desirable for a visual system because local estimates in a dense vector field provide the basis for the perception of 3D object motion - that is direction and speed of a moving object. This information is essential for segmenting objects from the background, for interpreting objects as well as for planning and executing actions in a dynamic environment.

If a single moving point, corner, or other unique feature serves as binocular input then intersection of constraint lines or triangulation in a binocular viewing geometry provides a straightforward and unique geometrical solution to the inverse problem. If, however, the moving stimulus has spatial extent, such as an oriented line or contour inside a circular aperture or receptive field then local motion direction of corresponding receptive fields in the left and right eye remains ambiguous, and additional constraints are needed to solve the inverse problem in 3D (Lages & Heron, 2010).

The inverse optics and the aperture problem are well-known problems in computational vision, especially in the context of stereo processing (Poggio, Torre, & Koch, 1985; Mayhew & Longuet-Higgins, 1982), structure from motion (Koenderink & van Doorn, 1991), and optic flow (Hildreth, 1984). Gradient constraint and related methods (e.g., Johnston et al., 1999) belong to the most widely used techniques of optic-flow computation based on image intensities. They can be divided into local area-based (Lucas & Kanade, 1981) and into more global optic flow methods (Horn & Schunck, 1981). Both techniques usually employ brightness constancy and smoothness constraints in the image to estimate velocity in an over-determined equation system. It is important to note that optical flow only provides a constraint in the direction of the image gradient, the normal component of the optical flow. As a consequence some form of regularization or smoothing is needed. Various algorithms have been developed implementing error minimization and regularization for 3D stereo-motion detection (e.g., Bruhn, Weickert & Schnörr, (2005); Spies, Jähne & Barron, 2002; Min & Sohn, 2006; Scharr & Küsters, 2002). These algorithms effectively extend processing principles of 2D optical flow to 3D scene flow (Vedula, et al., 2005; Carceroni & Kutulakos, 2002).

However, computational studies on 3D motion are usually concerned with fast and efficient encoding. Here we are less concerned with the efficiency or robustness of a particular algorithm and implementation. Instead we want to understand local and binocular constraints in order to explain characteristics of human 3D motion perception such as perceptual bias under uncertainty and motion estimation under ambiguity. Ambiguity of 2D motion direction is an important aspect of biologically plausible processing and has been extensively researched in the context of the 2D aperture problem (Wallach, 1935; Adelson & Movshon, 1982; Sung, Wojtach, & Purves, 2009) but there is a surprising lack of studies on the 3D aperture problem (Morgan & Castet, 1997) and perceived 3D motion.

The entire perceptual process may be understood as a form of statistical inference (Knill, Kersten & Yuille, 1996) and motion perception has been modeled as an inferential process for 2D object motion (Weiss, Simoncelli & Adelson, 2002) and 3D surfaces (Ji & Fermüller, 2006). Models of binocular 3D motion perception however are typically deterministic and predict only azimuth or change in depth (Regan & Gray, 2009). In the following we discuss probabilistic models of 3D motion perception that are based on velocity constraints and can explain perceptual bias under uncertainty as well as motion estimation under ambiguity.

The purpose of this chapter is to introduce and promote Bayesian models of binocular 3D motion perception because they provide local estimates of 3D velocity under uncertainty and ambiguity. We will focus on local constraints of motion perception using basic features such as single dots, lines or edges moving in 3D space because