The Scene in Front of You

Jan Koenderink
Utrecht University, Netherlands

Abstract

How it happens that you "see the scene in front of you" has been an unsolved problem for centuries. The modern main-stream account is essentially Marr's. Roughly speaking: "Vision is Inverse Optics". Although machine vision has made considerable progress along those lines, cracks in the foundation of this structure are clearly evident. Inverse optics yields ambiguous "solutions" (typically of infinite cardinality) and typically relies on assumptions that are rarely realistic (even physically impossible) and cannot be checked in the optical structure. Most importantly, it is fundamentally impossible to compute meaning. Where human vision is intentional (about the world), results of computations only acquire meaning in the mind of an end-user. Human vision is most certainly not "inverse optics", but more like "controlled hallucination", thus intentionality is there from the very start, not as the result of a computation. This is possible because human observers are both minds and (physical) machines. If machine vision is to emulate human capabilities (not necessary for many applications!) the Marrian, main-stream account has to be replaced with novel paradigms. I will discuss the structure of such systems, using various well known instances such as "Shape From Shading" as examples. I will also consider properties of human vision that might have potential interest for machine vision.

Syllabus: marr, inverse optics, intentionality, scenes, human observers, shape from shading, light fields, pictorial relief.