BEHAVIORAL VISUAL MOTION ANALYSIS

Yiannis Aloimonos

Computer Vision Laboratory
Center for Automation Research
Department of Computer Science
Institute for Advanced Computer Studies
University of Maryland
College Park, MD 20742

Extended Abstract

A new approach to addressing problems related to visual motion is proposed, namely the purposive approach. Instead of considering the various visual motion tasks as applications of the general structure from motion module, we consider them as independent problems and we directly seek solutions for them. As a result we can achieve unique and robust solutions without having to compute optic flow and without requiring a full reconstruction of the visual space, because it is not needed for the tasks. In the course of the exposition, we present novel solutions to various important visual tasks related to motion, such as the problems of motion detection by a moving observer, passive navigation, relative-depth computation, 3-D motion estimation, and visual interception, using as input only the spatial and temporal derivatives of the image intensity function. It turns out that the spatiotemporal derivatives of the image (i.e. the so-called normal flow) are sufficient to provide robust algorithms for the solution of many interesting visual tasks that do not require the full solution, but only part of it.

The support of ARPA, NSF (under a Presidential Young Investigator Award), ONR, Aliant Tech Systems, Inc., Texas Instruments, Inc., Sony and Westinghouse is gratefully acknowledged.

Thus, we do not have to solve the general structure from motion problem. The ability to create robust nontrivial behaviors suggests the possibility that visual perception could be studied as intelligent behavior. We point out some of the benefits and drawbacks of this paradigm that studies vision as a set of behaviors that recover the visible world partially, but well enough to carry out a task (purposive, animate or behavioral vision), we contrast it to the traditional paradigm of treating vision as a general recovery problem, and we present a formal framework formalizing behaviors and tasks.

This framework can be used for designing successful vision systems, by visually achieving various tasks without reconstruction but through the recognition of patterns, objects or situations. *What to recognize* is concerned with the questions we pose. The purposive paradigm calls for formulating questions that are directly related to visual tasks, i.e. that have a purpose. Knowledge of 3-D motion is much more than we need to answer the purposive question: Is this moving object coming closer to the observer? Purposive thinking leads us to pose questions whose answers will only help to solve the particular task at hand, and will not be of general use. This level of the paradigm is parallel to Marr's computational theory and insures that the resulting

algorithms will be of minimal complexity. *How to recognize* (patterns, objects or situations) is related to the algorithmic level of Marr's paradigm.

Qualitative vision calls for the development of algorithms that are simple, robust and based on qualitative techniques, such as comparisons of quantities or discrete classifications. Qualitative vision, which in the past has been wrongly called inexact, makes sense here because it is coupled with purposive vision, which formulates questions for which qualitative solutions are possible.