FINAL REPORT FOR ARO GRANT

Objectives

Our research is concerned with the theory and practice of visual information in a feedback loop, the underlying problem of controlled active vision. Controlled active vision requires the integration of techniques from control theory, signal processing, and computer vision. Visual tracking provides an important example of the need for controlled active vision. Tracking in the presence of a disturbance is a classical control issue, but because of the highly uncertain nature of the disturbance, this type of problem is very difficult. Visual tracking differs from standard tracking problems since the feedback signal is measured using imaging sensors. In particular, it has to be extracted via computer vision and image processing algorithms and interpreted by a reasoning algorithm before being used in the control loop. The development of robust, reliable visual tracking algorithms will be one of the main challenges in our proposed research.

In this research, we concentrate on the problem of tracking planar shapes which may be described by their closed boundaries, and the need for a new theory of observers. Observer design requires a dynamical model of the system to be observed and measurements that will make the observer (as a dynamical system) stable and robust with respect to unknown disturbances (e.g., measurement and system noise), and with respect to unmodelled dynamics. The type of observer used for a specific visual tracking problem is directly related to one's decision on how to (geometrically and dynamically) describe the object(s) to be tracked, leading for example to a description by means of coupled ordinary differential equations (in the finite-dimensional case) or partial differential equations (in the infinite-dimensional case). Accordingly, we need to develop a general theory of geometric observers which may be used in conjunction with geometric active contours for tracking.

Approach

As part of this program, we worked on the use of curvature-driven partial differential equations to explore problems in visual control and controlled active vision, and considered the sensor/estimation problem much more carefully in this framework. Many of the visual tracking techniques are based on concepts from optimal control (e.g., Hamilton-Jacobi theory) and so fit in very well with some of the modern robust and adaptive control strategies which explicitly take uncertainty into account. Filtering methods ranging from the classical Kalman filters valid for linear systems to the much more general particle filters also fit into this framework in a very natural manner. In particular, the particle filtering framework was applied to the space of continuous closed curves which is an infinite dimensional space. This is

a particularly difficult problem since generating Monte Carlo samples from a very large dimensional (theoretically infinite) system noise distribution is computationally complex. Moreover, the number of samples required for accurate filtering increases with the dimension of the system noise. In our approach, we approximate curve deformation using a time-varying finite dimensional representation. We can then formulate the problem as particle filtering with unknown static parameters and use a modification of a particle filter that has been shown to be asymptotically stable for tracking static parameters.

Scientific Barriers

Our work in the program was centered around the general area of using visual information in a feedback loop, the fundamental problem of controlled active vision. Controlled active vision requires the synergy of methodologies from control theory, signal processing, and computer vision. Visual tracking provides an important example of the need for controlled active vision. While tracking in the presence of a disturbance is a classical control problem, visual tracking raises new issues since the feedback signal is measured using imaging sensors. In particular, it has to be extracted via computer vision and image processing algorithms, and many times interpreted by a reasoning algorithm before being used in the control loop. In particular, for visual tracking one must integrate several basic vision tasks such as registration and segmentation into a real-time tracker. These problems are still not solved, yet in the past few years major progress has been made. Techniques such as geometric active contours fit very naturally into a control framework (they are derived from minimum energy principles from the calculus of variations), and they may be combined with Bayesian estimation techniques such as Kalman and particle filtering. Particle filters can be used in very general situations, can be used to track multiple objects, and have recently been incorporated into certain point-cloud registration algorithms. Much of our research is devoted to finding general principles for which trackers would function in the widest possible number of scenarios. Thus we must combine ideas from systems, vision, image processing, and statistical estimation theory.

Significance

The prevalence of biological vision in even very simple organisms, indicates its inherent utility for man-made devices. Cameras are rather simple, reliable passive sensing devices which are quite inexpensive per bit of data. Furthermore vision can offer information at a high rate with high resolution with a wide field of view and accuracy capturing multispectral information. Finally cameras can be used in a more active manner. Namely, one can include motorized

lenses and mounted on mobile platforms which can actively explore the surrounds and suitably adapt their sensing capabilities. Computer vision has formulated several approaches for interpreting the signals, opening the possibility that control laws can be based on more abstracted descriptions. These problems all become manifest when one attempts to use a visual sensor in an uncertain environment, and to feed back in some manner the information. These issues represent some of the main challenges for our research in controlled active vision. Control is very powerful in treating uncertainty, and with the newer partial differential methodologies in computer vision, there is a strong mathematical and systems-theoretic fit. The collaborative effort joining the efforts and talents and the systems and vision communities has recently been treating in a "first principles" manner some of the major problems remaining in active vision, and thus creating more reliable robust systems in which the primary sensor is based largely on some imaging device.

Accomplishments

(1) ***Robust Tracking and Target Reacquisition on 3D Range Data with Particle Filtering and Online Shape Learning:*** We have developed an algorithm for tracking an object of interest based on 3D range data. We employ particle filtering and active contours to estimate the global motion of an object and its local deformations, respectively. A region-based active contour model driven by the Bhattacharyya gradient flow is adopted for its robustness to noise in a cluttered environment. The proposed algorithms take advantage of both range information and contour-tracking while at the same time dealing with the challenging situation that occurs when objects being tracked disappear from the image domain and reappear later. To cope with this problem, a PCA-based shape analysis is proposed, which defines a shape similarity energy to find target candidates that are as close as possible to the template shape obtained online from the previous segmentations, rather than the prior training shapes. Experimental results show the robustness and practicality of the proposed algorithm for real tracking problems.

(2) ***Point-Cloud Registration:*** We have developed a particle filtering approach for the problem of registering two point sets that differ by a rigid body transformation. Typically, registration algorithms compute the transformation parameters by maximizing a metric given an estimate of the correspondence between points across the two sets of interest. This can be viewed as a posterior estimation problem, in which the corresponding distribution can naturally be estimated using a particle filter. In this work, we treat motion as a local variation in pose parameters obtained by running a few iterations of a certain local optimizer. Employing this idea, we introduce stochastic motion dynamics to widen the narrow band of convergence often found in local optimizer approaches for registration. Thus, the novelty of our method is threefold: Firstly, we employ a particle filtering scheme to drive the point set registration process. Secondly, we present a local optimizer that is motivated

by the correlation measure. Thirdly, we increase the robustness of the registration performance by introducing a dynamic model of uncertainty for the transformation parameters. In contrast with other techniques, our approach requires no annealing schedule, which results in a reduction in computational complexity (with respect to particle size) as well as maintains the temporal coherency of the state (no loss of information). Also, unlike some alternative approaches for point set registration, we make no geometric assumptions on the two data sets. Experimental results are provided that demonstrate the robustness of the algorithm to initialization, noise, missing structures or differing point densities in each sets, on challenging 2D and 3D registration tasks.

(3) ***Fast Level Set Methodss:*** The level set method for curve evolution is a popular technique used in image processing applications. However, the numerics involved make its use in high performance systems computationally prohibitive. This paper proposes an approximate level set scheme that removes much of the computational burden while maintaining accuracy. Abandoning a floating point representation for the signed distance function, we use the integral values to represent the interior, zero level set, and exterior. We detail rules governing the evolution and maintenance of these three regions. Arbitrary energies can be implemented with the definition of hree operations: initialize iteration, move points in, move points out. This scheme has several nice properties. First, computations are only performed along the zero level set. Second, this approximate distance function representation requires only a few simple integer comparisons for maintenance. Third, smoothness regularization involves only a few integer calculations and may be handled apart from the energy itself. Fourth, the zero level set is represented exactly removing the need for interpolation off the interface. Lastly, evolution proceeds on the order of milliseconds per iteration using conventional uniprocessor workstations. To highlight its accuracy, flexibility and speed, we have demonstrated the technique on standard intensity tracking and stand alone segmentation.

Conclusions

The tracking of multiple objects in complex adversarial environments using a vision sensor is certainly a problem of major importance. Typical approaches to problems such as segmentation and registration many times fail in realistic challenging environments in which there is significant clutter, occlusions, and noise. In our ARO sponsored work, we have been developing techniques in control, estimation, and computer vision for real-time visual tracking applications. In addition to segmentation problems because of the temporal nature of the target data, filtering theory (e.g., Kalman and nonlinear filtering) becomes very relevant. Geometric active contours incorporating statistical and geometric information that may naturally be employed in a Bayesian filtering framework are key tools in our research efforts.