Inertially Assisted Stereo Tracking for an Outdoor Rover

Kevin Nickels
knickels@engr.trinity.edu
Engineering Science
Trinity University
San Antonio, TX 78212

Eric Huber
huber@roboteyes.com
Metrica, Inc.
Robotics and Automation Group
NASA Johnson Space Center
Houston, TX 77058

Abstract

Ego-motion (self-motion) of the camera is a considerable problem in outdoor rover applications. Stereo tracking of a single moving target is a difficult problem that becomes even more challenging when rough terrain causes significant and high-acceleration motion of the camera in the world.

This paper discusses the use of inertial measurements to estimate camera ego-motion and the use of these estimates to augment stereo tracking. Some initial results from an outdoor rover will be presented, illustrating the efficacy of the method. The method causes fast but predictable image location transients, but reduces the amplitude of image location transients due to the rough terrain.

1 Introduction

The Robotic Systems Technology Branch of the National Aeronautics and Space Administration (NASA) is currently working on the development of an Extra-Vehicular Activity Robotic Assistant (ERA) under the sponsorship of the Surface Systems Thrust of the NASA Cross Enterprise Technology Development Program (CETDP). This will be a mobile robot that can follow a field geologist during planetary surface exploration, carry his tools and the samples that he collects, and provide video coverage of his activity.

Prior experiments have shown that for such a robot to be useful it must be able to follow the geologist at walking speed over any terrain of interest. Geologically interesting terrain tends to be rough rather than smooth. The commercial mobile robot that was recently purchased as an initial testbed for the ERA project, an ATRV Jr., is capable of faster than walking speed outside but has no suspension. Its wheels with inflated rubber tires are attached to axles that are connected directly to the robot body. Any angular motion of the robot produced by driving over rough terrain will directly affect the pointing of the on-board stereo cameras. The resulting image motion makes tracking of the geologist more difficult. This either requires the tracker to search a larger part of the image to find the target from frame to frame or to search mechanically in pan and tilt whenever the image motion is large enough to put the target outside of the image in the next frame.

This paper explains the design and implementation of a set of Kalman filters that utilize the output of the angular rate sensors and linear accelerometers on the robot to estimate the motion of the robot base. Section 2 reviews the hardware utilized in this project, and Section 3 describes the Kalman filters developed to estimate this motion.

The motion of the stereo camera pair mounted on the robot that results from this motion as the robot drives over rough terrain is then straightforward to compute. Section 4.1 describes the computation of camera motion from base motion. This computation is similar to that done for image stabilization on commercial video cameras [6] [2] but we do not, in general, wish to cancel camera motion but rather to combine it with other information. Also unlike video cameras, we do not currently utilize motion vectors derived from images as in general the cameras are undergoing commanded motions to track a target.

Image motion estimates based on inertial data may be used, for example, to command the robot's onboard pan-tilt unit to compensate for the camera motion induced by the base movement. This has been accomplished in two ways: first, the standalone head stabilizer described in Section 4.2 has been implemented and second, the estimates have been used to influence the search algorithm of the stereo tracking algorithm as described in Section 4.3. Studies of the image motion of a tracked object, shown in Section 5, indicate that the image motion of objects is suppressed while the robot is crossing rough terrain.

2 Rover Hardware

The rover is a modified ATRV Jr., from RWI. The wheels have been mounted on rigid extensions to provide adequate ground clearance, and a tower has been

added to the top for the stereo vision hardware. Figure 1 shows a cartoon of the rover. The rover comes
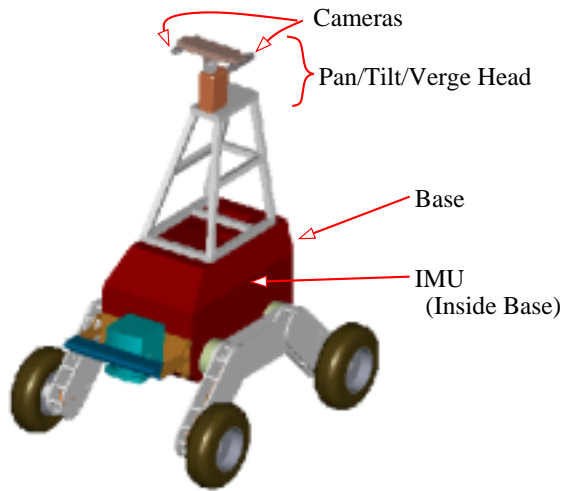

Figure 1: The EVA Robotic Assistant testbed

equipped with a magnetic compass, an Inertial Measurement Unit (IMU), and two on-board computers. It has been augmented to include a pan-tilt-verge (PTV) head and two cameras with framegrabbers. The relevant components are described briefly below. Mathematical models for these sensors are given later in the paper.

2.1 Sensors Used

The sensors considered in this paper are three mutually orthogonal gyrometers, three mutually orthogonal linear accelerometers, and a magnetic compass.

A gyrometer measures angular velocity about a single axis. These measurements are assumed to be corrupted by gyro biases [3]. These biases are commonly estimated for purposes of compensation (see below for mathematical sensor models used.) After compensation, the angular velocities recovered can be integrated to arrive at an estimate of the rotation of a body relative to some fixed initial orientation.

A linear accelerometer measures acceleration along a single axis. The accelerations can be integrated to arrive at linear velocities, and integrated again to arrive at position relative to some initial position. In this work, the accelerometers were not used in this fashion, but were used to measure the direction of the gravitational vector while the rover was at rest. See [8] for more discussion on inertial data.

The linear accelerometers and gyros used in this project were packaged in a single Inertial Measurement Unit (IMU), the DMU-6X from Crossbow. A

magnetic compass yields a bearing with respect to magnetic north. The magnetic compass used in this project was the TCM2 from Precision Navigation.

2.2 Actuators Used

The stereo pan-tilt-verge (PTV) heads utilized in this paper are the Zebra Vergence from Pyxis Corp (formerly Helpmate, formerly TRC) and the Biclops from Metrica. Each of these heads accepts movement commands via a serial port from an external computer. Each head supports two cameras that are used for image acquisition. Components based upon the Common Object Request Broker Architecture (CORBA) [4] have been developed for image acquisition and control of the heads. Wasson, Kortenkamp and Huber [9] describe the interaction of components to control gaze.

3 Kalman filtering

This work utilizes six one-dimensional Kalman filters (or equivalently, one six-dimensional Kalman filter with a diagonal system covariance.) The six quantities estimated are the three drifts associated with the three gyros and the three roll-pitch-yaw angles ϕ, ψ, and θ. Orientation given by roll-pitch-yaw angles is defined by taking a base coordinate system (x_0, y_0, z_0), rotating about the x_0 axis by the roll angle ϕ, rotating about the y_0 axis by the pitch angle ψ, and finally rotating about the z_0 axis by the yaw angle θ. See [7] for more discussion on the representation of rotation.

We distinguish between two distinct movement modes of the robot: rest mode (RM) and maneuvering mode (MM). If the rover is at rest we exploit some useful measurements and assumptions. If the rover is maneuvering, these assumptions are not made, and the general form of the filters are used. In this way, we can maximally exploit the information provided by the various sensors.

3.1 Determination of Mode

One important characteristic of the developed filters is the difference in behavior when the rover is at rest and when the rover is maneuvering. Assumptions about the inertial sensors, described below, change based upon the mode. Therefore, a high-quality estimate of the mode of the rover is required.

We have designed a transient detector to distinguish between these modes of operation. This detector utilizes hysteresis to handle outliers. Each gyrometer data point is compared against a running average of the previous 30 samples. If the angular velocity is greater than 0.5 degrees/s from this average, that

data point is defined to be a transient data point. If 5 consecutive data points are labelled as transient data points, the mode is defined to be MM. Leaving MM should be a more conservative transition, so 30 consecutive non-transient data points are required to leave MM and enter RM. No mode changes are made until these thresholds are reached. All of these thresholds have been defined experimentally and are tunable. This detector allows robust determination of the motion state of the rover.

3.2 Gyro Drift Estimates - RM

If the robot is at rest, the measured angular velocities consist completely of drift. In this case, simple one-dimensional discrete Kalman Filters are used to estimate drift about each axis. The assumed models are

$$d_\theta(k+1) = d_\theta(k) + w(k) \qquad \text{(Motion Model)}$$

$$\dot{\theta}_g(k) = \dot{\theta}(k) + d_\theta(k) + v(k) = d_\theta(k) + v(k)$$
$$\text{(Measurement Model)}$$

where $d_\theta(k)$ is the drift for the gyro measuring angle θ at sample k, and the measured angular velocity $\dot{\theta}_g(k)$ consists entirely of drift and noise. The noise sequences $w(k)$ and $v(k)$ are assumed to be zero-mean Gaussian white noise of covariance $Q(k)$ and $R(k)$, respectively. This leads to a Kalman filter implementation of

$$K_\theta(k) = P_\theta(k-1)(P_\theta(k-1) + R(k))^{-1}$$
$$\text{(Kalman Gain)}$$

$$\widehat{d}_\theta(k) = \widehat{d}_\theta(k-1) + K_\theta(k)(\theta_g(k) - \widehat{d}_\theta(k-1))$$
$$\text{(Estimate Update)}$$

$$P_\theta(k) = (I - K_\theta(k))P_\theta(k-1) + Q(k)$$
$$\text{(Uncertainty Update)}$$

for computing the estimate $\widehat{d}_\theta(k)$ of the drift on the θ axis. The computation of the Kalman gain matrix $K_\theta(k)$ and the state estimate uncertainty $P_\theta(k)$ are straightforward. See [1] for more explanation on the derivation of Kalman filters. The filters for the ϕ and ψ drifts are analogous.

3.3 Attitude Estimate - RM

At rest, the attitude of the robot can be estimated based on the projection of gravity (which is assumed to be directed along the $+z$ axis of the inertial frame)[8],

$$\psi(t) = -sin^{-1}g_x/g$$
$$\phi(t) = sin^{-1}g_y/g \ cos(\psi(t))$$

These measurements of attitude are combined with previous estimates of attitude in one-dimensional Kalman filters to achieve smoothing and outlier rejection.

3.4 Gyro Drift Estimates - MM

If the ERA is maneuvering, the simplifying assumptions made in the previous sections are invalid. Therefore, we use different motion models for this mode. We simply maintain a constant drift estimate and increase the uncertainty of the estimate with time. In essence, we are neglecting the observation by setting the measurement covariance $R(k) = \infty$, yielding a Kalman gain of zero.

$$\widehat{d}_\theta(k) = \widehat{d}_\theta(k-1) \quad \text{(No Estimate Update)}$$
$$P_\theta(k) = P_\theta(k-1) + Q(k) \quad \text{(Uncertainty Update)}$$

3.5 Angular Velocity Estimates - MM

To estimate the actual angular velocities in this case, we subtract the gyro drift estimates from the gyrometer reading:

$$\widehat{\dot{\phi}}(k) = \dot{\phi}_g(k) - d_\phi(k)$$
$$\widehat{\dot{\psi}}(k) = \dot{\psi}_g(k) - d_\psi(k)$$
$$\widehat{\dot{\theta}}(k) = \dot{\theta}_g(k) - d_\theta(k)$$

3.6 Attitude Estimation - MM

As the vehicle acceleration is superimposed on the gravitational acceleration, the attitude estimates during maneuvering are instead derived from integration of the angular velocities, corrected by the drift estimates as described above.

$$\phi_g(t) = \phi_g(t - \Delta t) + \int_{t-\Delta t}^{t} \left[\dot{\phi}_g(t) - \widehat{d}_\phi(t) \right] dt$$

$$\psi_g(t) = \psi_g(t - \Delta t) + \int_{t-\Delta t}^{t} \left[\dot{\psi}_g(t) - \widehat{d}_\psi(t) \right] dt$$

$$\theta_g(t) = \theta_g(t - \Delta t) + \int_{t-\Delta t}^{t} \left[\dot{\theta}_g(t) - \widehat{d}_\theta(t) \right] dt$$

where Δ is the sampling period. This integration can be done via any numerically sound (e.g. rectangular, trapezoidal, Runge-Kutta) algorithm. Trapezoidal integration is used in our implementation. These estimates are also folded into one-dimensional Kalman filters to achieve smoothing and outlier rejection.

4 Use of Base Motion Estimates

Estimates of angular velocity and angular position (net rotation since initialization) can be used to correct

for *ego-motion* in images. Ego-motion is defined to be image motion due to the movement of the camera in the world. This is distinct from object motion, which is image motion caused by the movement of the tracked object in the world. Object motion is not addressed in this paper.

The following sections describe how we translate base rotation to camera motion, and two methods for utilizing these estimates. In Section 4.2, we describe a standalone stabilization behavior that can be combined with manual control to provide a stabilized image. In Section 4.3, we describe the use of these estimates in a stereo tracking application.

4.1   Transforms Used

Given the base orientation and current PTV configuration (pan, tilt, and verge angles), a homogeneous transform (see, e.g. [7]) can be created that relates an object of interest in the world coordinate system to that object in a camera coordinate system:

$$P_{cc} = H_{cc}^{base} H_{base}^{world} P_{world}$$
$$H_{cc}^{base} = Rot(x, tilt) \times$$
$$Rot(z, pan \pm verge/2) \times H_{base}^{mount}$$
$$H_{base}^{world} = Rot(z, \theta) \times Rot(y, \theta) \times Rot(x, \phi)$$

where H_a^b is the homogeneous transform from coordinate system a to coordinate system b, P_a are the coordinates of a point in coordinate system a, and $Rot(a, \theta)$ indicates a rotation of θ about the a axis.

Given this transform and the world position of an object, the position of the object in the current camera coordinate frame can be determined. This position can be then used to command the cameras to center the object in the field of view, as described in the following section. The transform can also be passed to object tracking software to initiate the object search in the new image, as described in Section 4.3.

4.2   Standalone Stabilization

In a "standalone" configuration, there is a single point in the world coordinate system that we wish to *stabilize*. We define *stabilization* as the act of maintaining the image of a particular object in a fixed location in the camera image. Note that as the actuators used (the Zebra Vergence and Biclops PTV heads) have no degree of freedom to compensate for roll, only a single point can be stabilized. The remainder of the target object will rotate in the image with base roll. The location of this point is assumed to change slowly relative to frame rate, making this configuration useful, for example, for manual control of the view.

Once the transforms described in Section 4.1 are computed, a command is sent to the PTV head to send the object of interest to a fixed point in the camera coordinate system (for example, directly centered in front of the camera). Both PTV heads utilized in this experiment (see Section 2.2) contain sophisticated control algorithms in either the hardware or the CORBA drivers, ensuring smooth trajectories without further attention to control strategy.

4.3   Interaction with Stereo Tracking

Stereo tracking software has been developed at Johnson Space Center (JSC) that will track the torso and arm of a suited astronaut [5]. One objective of our work is to utilize the ego-motion estimates in this tracking software. This section describes this process.

In each new frame, the tracking software begins searching an image for several textural features which make up the target. The initial image location of this search has been the previous image location of the target. In situations where significant ego-motion is encountered, the performance of the stereo tracking algorithm can be effected. The amount of the image that is searched for the target may need to be increased, thus slowing the algorithm and increasing the likelihood of false alarms. If the target is lost due to the target leaving the search area or field of view, the rover may need to stop while the target is reacquired.

If ego-motion estimates are available, as described above, the initial image location for search can instead be set to be the *expected* image location of the target. This location is defined as the *new* image location of the *previous* world location of the target. This should allow the reduction of the search area and a subsequent increase in the frame rate of the object tracking software.

In some cases, the target may actually leave the field of view of the camera. Purely visual tracking is at a significant disadvantage here. If the apparent motion of the target has been consistent for several frames, the visual motion may be sufficient to reacquire the target after re-servoing the head. However, it is much more likely that in this situation either the target, the rover, or both have changed acceleration and velocity. An algorithm utilizing purely visual information will fail in this situation. If the target leaves the field of view, the stereo tracker can utilize the expected image location of the object (which will be outside the viewable portion of the image plane) to direct the PTV head to servo in the appropriate direction and hopefully reacquire the target.

In the event that the target is completely lost, inertial data is also useful. In this case, a volume of space

parallel to the ground plane is systematically searched for the target (the torso and head of a suited astronaut). If inertial data is not used, this search may not occur parallel to the ground plane, and will not be as likely to succeed.

In each of these ways, inertial data is used to augment existing JSC stereo tracking software. With these augmentations, the effect of rough terrain on stereo tracking performance is reduced.

5 Results

Figures 2 and 3 illustrate the performance of the filters for the gyro drifts and orientation angles. Figure 4 illustrates the behavior of the standalone stabilization.
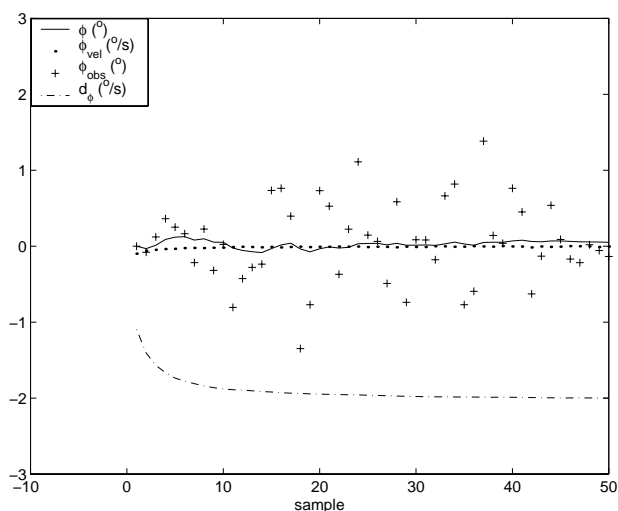

Figure 2: Initialization of drift estimate (roll only)

Figure 2 shows the initialization of the drift estimate d_ϕ while the base is at rest. The high initial uncertainty of the drift estimate causes the measurements ϕ_{obs} to significantly affect the estimate. After the drift estimates become more certain, new measurements begin to affect the drift estimates less. The velocity ϕ_{vel} quickly approaches zero.

Figure 3 illustrates both sets of Kalman filters: the orientation angles and drift estimates. After many samples, the estimate for the drift is fairly certain. Upon entering an angular transient, there is a detection lag of several samples. Therefore, during this time (near sample 635) the drift estimate does not change appreciably. As described in Section 3, once a transient is detected, updates to the drift estimates are suspended for the duration of the transient. In this
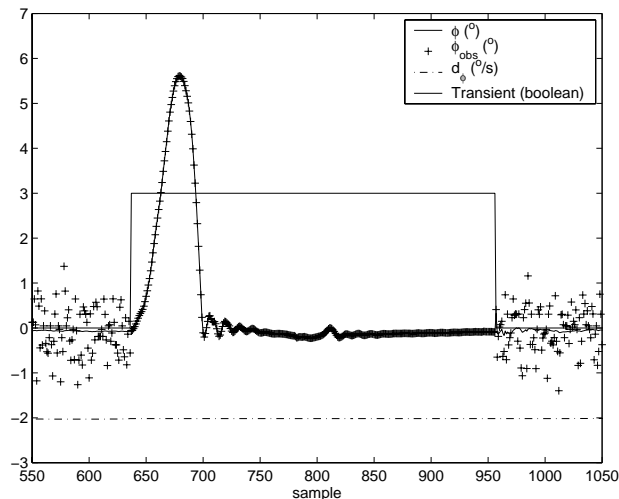

Figure 3: Drift estimation during transient (roll only)

test, the robot begins at rest, drives over an obstacle, then is at rest again.

This figure also shows the difference in uncertainty between the at-rest observations of the attitude (derived from the accelerometers) and the observations derived by integrating the gyro measurements. The derived measurements are more precise, but are subject to a slow drift over time, while the accelerometer-derived measurements are bias-free, but have a high degree of uncertainty. Both types of measurements are used over time, yielding the behavior shown in Figure 3: a filter that responds quickly and accurately to measure transient behavior, but will reset the attitude estimates any time the base is at rest.

Figure 4 shows the location in image coordinates (u along the horizontal direction, v along the vertical) of an object of interest. This object drifts slowly lower in the image as the rover moves forward toward the object (it is located slightly below the PTV head). The approximate ground truth location of the object without terrain is shown with dashed lines. The hand-tracked image location of the object with fixed gaze (constant pan, tilt, and verge angles on the PTV head) is shown with dotted lines. As can be seen from the figure, with fixed gaze there is a large vertical transient near samples 50-100. This corresponds to the front tire of the rover encountering an obstacle. The other transient (near samples 190-220) corresponds to the rear tire encountering the same obstacle. With stabilization turned on (the dot-dashed lines,) the algorithm described in Section 4.2 is controlling the PTV head, and the amplitude of both transients are smaller.

There is a tradeoff for the stabilization, however. As can be seen from Figure 4, stabilization currently
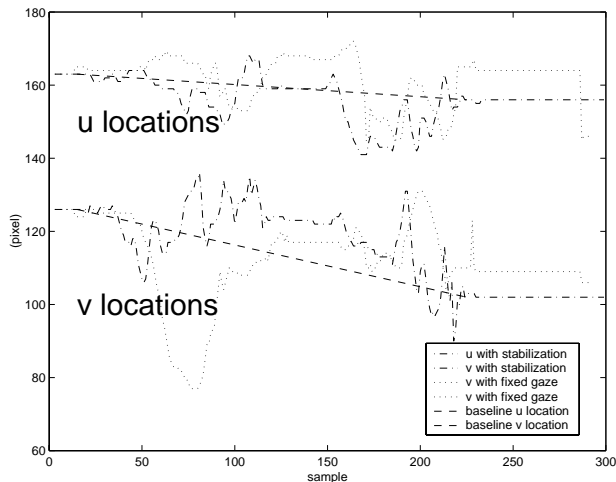
3082

Figure 4: Stabilization during terrain traversal

induces a low-frequency vibration in image location. The object of interest remains roughly in the center of the center of the image, as intended, but the picture appears to shake. We believe that this effect is due to the point-to-point move commands currently used to command the PTV heads. We are working to integrate the ego-motion estimation with existing tracking software that drives the PTV heads by updating desired angular velocities of the PTV head, which may eliminate or reduce this effect. Note that even with the artifacts, the amplitude of the transients is decreased.

6 Future Work

The primary extension to this work will be to complete the integration of the ego-motion estimates generated by this filter with the existing JSC stereo tracking software and to evaluate the efficacy of this upgrade. Evaluation of the behavior of the angular estimates in the field may take place in September 2000, during the scheduled tests in Arizona.

Less immediate extensions include the elimination of the explicit notion of *modes* of operation to be replaced by a continuous scale that can be used to smoothly transition between exploiting rest-mode assumptions and the general form of the filter. Feedback on actual camera motion could be generated by the stereo vision software and incorporated into the attitude estimates.

Finally, positional estimates have not been addressed at this point. Positional information from a GPS reciever, wheel encoders, and the linear accelerometers could be combined to arrive at estimates for the position, linear velocity, and linear acceleration of the robot. This information could be usedto generate the three-dimensional path of the rover.

7 Conclusions

Inertial data can be used to compensate for ego-motion in images taken from an outdoor rover. This compensation can be treated as a standalone behavior, to keep a specified object of interest centered in an image, or as an input to a more complex object tracking algorithm. Initial tests reveal that some low-frequency oscillation was introduced as a result of the stabilization, but that the amplitude of image location transients due to obstacles in the path of an outdoor rover decreased. This should expand the range of speed and surface roughness over which the rover should be able to visually track and follow a field geologist.

References

[1] R. G. Brown. *Introduction to Random Signal Analysis and Kalman Filtering*. John Wiley & Sons, New York, 1983.

[2] A. Engelsberg and G. Schmidt. A comparative review of digital image stabilising algorithms for mobile video communications. *IEEE Transactions on Consumer Electronics*, 45(3):591–597, 1999.

[3] E. Foxlin. Inertial head-tracker sensor fusion by a complementary separate-bias Kalman filter. In *Proc. 1996 Virtual Reality Annual International Symposium*, 1996.

[4] M. Henning and S. Vinoski. *Advanced CORBA Programming with C++*. Addison-Wesley-Longman Inc, 1999.

[5] D. Kortenkamp, E. Huber, and R. Bonasso. Recognizing and interpreting gestures on a mobile robot. In *Proceedings of the National Conference on Artificial Intelligence*, volume 2, pages 915–921, August 1996.

[6] B. Schweber. Image stabilization shows diversity of engineering approaches. *EDN*, page 79, October 26 2000.

[7] M. Spong and M. Vidyasagar. *Robot Dynamics and Control*. John Wiley & Sons, New York, 1989.

[8] J. Vaganay, M. J. Aldon, and A. Fournier. Mobile robot attitude estimation by fusion of inertial data. In *Proc. IEEE International Conference Robotics and Automation*, pages 277–282, 1993.

[9] G. Wasson, D. Kortenkamp, and E. Huber. Integrating active perception with an autonomous robot architecture. *Robotics and Automation Journal*, 29:175–186, 1999.