Intent Recognition for Human-Robot Interaction

Andreas G. Hofmann and Brian C. Williams

Computer Science and Artificial Intelligence Lab, MIT
32 Vassar St. rm. 32-275
Cambridge, MA 02139
hofma@csail.mit.edu, williams@mit.edu

Abstract

Effective human-robot cooperation requires robotic devices that understand human goals and intentions. We frame the problem of intent recognition as one of tracking and predicting human actions within the context of plan task sequences. A hybrid mode estimation approach, which estimates both discrete operating modes and continuous state, is used to accomplish this tracking based on possibly noisy sensor input. The operating modes correspond to plan tasks, hence, the ability to estimate and predict these provides a prediction of human actions and associated needs in the plan context. The discrete and continuous estimates interact in that the discrete mode selects continous dynamic models used in the continuous estimation, and the continuous state is used to evaluate guard conditions for mode transitions. Two applications: active prosthetic devices, and cooperative assembly, are described.

Introduction

Augmenting human capabilities with automated, cooperative robotic devices is important for a wide variety of tasks, including, construction, assembly, repair, search and rescue, and general assistance with every-day tasks for the elderly and handicapped. Effective human-robot cooperation requires robotic devices that understand human goals and intentions. These devices must have the capability to track and predict human intention and motion within the context of overall plan task sequences, based on a variety of sensor inputs. Such a capability will enable a fundamentally new kind of collaboration between humans and machines; one where the robots' actions are based, primarily, on implicit rather than explicit commands from humans.

We are currently engaged in two projects of this nature. The first deals with active prosthetic devices [Blaya and Herr, 2004; Herr and Wilkenfeld, 2003], where, in contrast to currently existing passive devices, actuators are able to exert force and move parts of the devices. The actuators must be carefully controlled, based on a recognition of what the human user is trying to accomplish. The second project involves recognition of human gestures and actions for assembly, in order to support cooperative assembly tasks.

Active Ankle Prosthetic

For the prosthetic application, we seek to appropriately control an active ankle prosthetic, shown in Fig. 1, based on recognition of the type of walking task the user is engaged in. The prosthetic has a single actuator at the ankle, which is used to adjust the pitch angle of the foot with respect to the shank. This capability improves the user's ability to perform a variety of walking tasks, including walking up and down stairs, walking up and down slopes, as well as level ground walking, in a more natural, safe, and efficient manner. For example, when walking down stairs, the pitch angle is greater than during level ground walking. This results in the toes touching before the heel as the foot descends to the next step, allowing for better absorption of impact forces. This is in contrast to level ground walking, where the heel strikes before the toe.

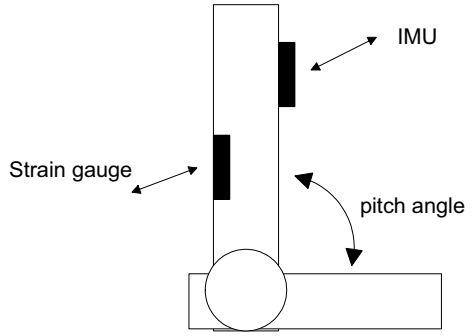

Fig. 1 – Active ankle prosthetic – an IMU attached at the shank measures inclination and linear acceleration; a strain gauge on the shank measures contact forces.

In order for the pitch angle control to be safe, the system must recognize user intent in a timely manner. For example, the system should recognize a transition from level ground walking to walking down stairs soon enough for the pitch angle to be adjusted before the first descending step. Intent recognition is based on sensor information from two sources: an Inertial Measurement Unit (IMU), as well as a strain guage, both of which are mounted on the shank of the prosthetic.

The IMU provides very accurate information about the three-dimensional orientation of the shank. It also provides translational acceleration, but this has some error. This acceleration error can cause drifting of velocity and position estimates. The strain gauge is used to determine if the foot is on the ground or not.

We represent the state of the combined user/prosthetic system using a discrete/continuous hybrid state vector. The type of task being performed (taking a step on level ground, taking a step to walk down stairs, etc.) is represented using a discrete mode variable, and position and velocity state is represented by continuous variables. We estimate and predict this hybrid state using a hybrid mode estimation architecture [Williams et al., 2001, Hofbaur and Williams, 2002] that combines the predictive capabilities of dynamic models, with observations from the sensors on the prosthetic.

We frame this tracking and prediction process as belief state update for a *hybrid Hidden Markov Model* (hybrid HMM). In a hybrid HMM, each discrete mode has an associated continuous dynamics for the continuous state variables. The continuous state variables and system observations are given by stochastic difference equations. Mode transition is a probabilistic function of the current mode and continous state estimates. We use a Hybrid Markov Observer (Fig. 2) to interpret the hybrid HMM.

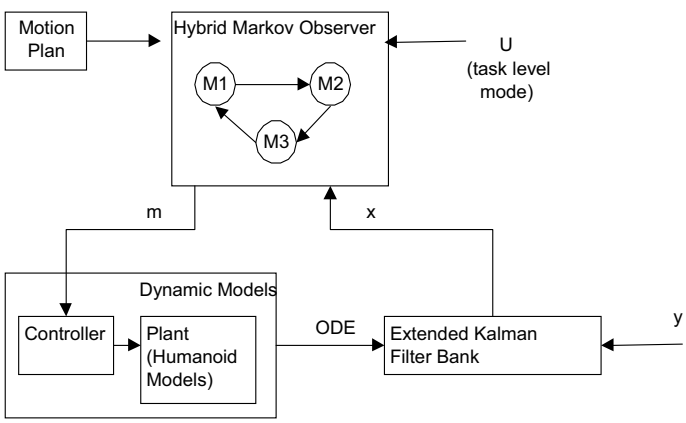

Fig. 2 – Hybrid Markov Observer – The hybrid markov observer interprets the hybrid HMM, using an extended Kalman filter bank to estimate continuous state from sensor information. Filters in the filter bank are selected based on discrete mode.

The observer computes a sequence of hybrid state estimates, each of which is a tuple $\hat{\mathbf{x}}_k = \langle \hat{\mathbf{x}}_{d,k}, p_{c,k} \rangle$, where $\hat{\mathbf{x}}_{d,k}$ is the estimate of the discrete mode, and $p_{c,k}$ is the continuous state estimate expressed as a multi-variate probability distribution function with mean $\hat{\mathbf{x}}_{c,k}$ and covariance matrix $\mathbf{P}_k$.

A key difference between this hybrid estimation approach and standard HMM belief-state update is that hybrid estimation tracks a set of trajectories, rather than aggregating trajectories with the same mode [Hofbaur and Williams, 2002]. This difference is crucial in that it allows for correct utilization of the continuous state estimates provided by the Kalman filter bank. However, it can lead to a combinatoric explosion of possible trajectories being tracked. To solve this problem, a k-best filtering approach is used, where only the trajectories with the k best belief states are tracked.

Intent Recognition for Cooperative Assembly

Our second application area involves intent recognition for the purpose of assisting in an assembly task. For example, in order for a robot to assist with a task like assembling a piece of furniture, it must be able to not only recognize basic movements like grasping or lifting, but it must also be able to place these in the context of an overall plan [Williams et al., 2003]. This would allow the robotic assistant to anticipate the needs of the human. For example, by knowing that the next step in the assembly plan involves joining two parts, the robotic assistant could be ready with the parts, holding them together, so that the human can put screws into place.

For this application, we use a hybrid estimation approach, similar to the one for the ankle prosthetic. In this application, the discrete mode represents contextual information about the overall plan being executed. The HMM representation is generated automatically from a high-level plan specification language called RMPL [Williams, 2003].

References

Blaya, J., Herr, H., “Adaptive Control of a Variable-Impedance Ankle-Foot Orthosis to Assist Drop-Foot Gait. IEEE Transactions on Neural Systems and Rehabilitation Engineering 2004; 12(1): 24-31.

Herr, H., Wilkenfeld A., “User-Adaptive Control of a Magnetorheological Prosthetic Knee”, Industrial Robot: An International Journal 2003; 30: 42-55.

Hofbaur, M., Williams, B., “Mode Estimation of Probabilistic Hybrid Systems”, HSSC 2002, LNCS 2289, pp. 253-266, Tomlin and Greenstreet (Eds.), Springer-Verlag

Williams, B., Chung, S., Gupta, V., 2001. "Mode Estimation of Model-based Programs: Monitoring Systems with Complex Behavior." *Proceedings of the International Joint Conference on Artificial Intelligence*, Seattle, Wa.

Williams, B., Ingham, M., Chung, S., and Elliott, P. (2003), Model-based Programming of Intelligent Embedded Systems and Robotic Space Explorers, invited paper in *Proceedings of the IEEE*: Special Issue on Modeling and Design of Embedded Software, vol. 9, no. 1, pp. 212-237.