Motor Control Programming through Demonstration

Mike Wessler

Artificial Intelligence Laboratory
Massachusetts Institue Of Technology
Cambridge, Massachusetts 02139

http://www.ai.mit.edu

The Problem: Develop a technique for programming a simulated biped robot to walk, without requiring a-priori knowledge of its forward or reverse kinematics. The programmer presents demonstration "films" of walking behavior to an online learning system, which attempts to mimic the motions of walking. The programmer can assist the online learning by suggesting gross movement techniques in trouble spots. The system then refines these into a more graceful walk.

Motivation: Robots in the Leg Lab are typically programmed using *Virtual Model Control*[1], a technique for developing motor control using high level, intuitive controllers instead of low level torque control. While the technique works very well in most cases, it is brittle to changes or mismeasurements of the kinematics.

Instead, we would like to be able to program a robot by presenting a demonstration of the task goal. The learning system will use the demonstration as a starting point in online learning. The same principles that allow this system to learn the motor control will allow it to adapt to small changes in the structure or mass of the robot. Parameters that need to be tuned by hand using current techniques will be automatically adapted by this algorithm.

Previous Work: There is quite a lot of previous work in bipedal locomotion in addition to the work done here in the Leg Lab. The Honda P3 robot plays back modified recordings of a human being on all its joints except the ankles, which are used for balance.[2] The recordings are tweaked offline to guarantee that the robot does not fall. Using prerecorded data, the robot displays extremely life-like walks. However, a recording, once started, must be played through to the end. At the other extreme, Miller[3] uses a very simple, general algorithm to generate a walking gait, and employs several neural networks to tune the few parameters necessary to keep balance.

Schaal and Atkison[4] have demonstrated the feasibility of using demonstration to shorten the learning time in motor control applications. They note that generating a policy based on demonstration dramatically improves the initial success of reinforcement learning. Much of their success in teaching a robot to juggle[5] comes from a clever translation of the problem from continuous motion to periodic, discrete events.

Approach: The learning system is broken into two parts. The first, offline part is given a recording of a walk cycle, a sequence of joint angles and positions. It breaks the continuous motion into a suggested sequence of gross motor movements. These movements are loosely based on the types of motor control seen in some human motion: ballistic launching (the start of a swinging motion), braking (the end of a ballistic swinging motion), and balance (inverted pendulum style maintenance of some parameter).

The online portion of the algorithm uses the gross motor sequence as a template and the recording of the walk cycle as a guideline and critic to learn to walk. The system adapts both the parameters of each type of motion and their trigger points. The system has fairly rapid reward feedback by tracking its motion against the motion in the recording. There is no need to wait for the entire robot to fall over before deciding that something is wrong. With only a single walk cycle as input data, however, there is no demonstration of how to recover from deviations. The user can note places where the algorithm fails and make recovery suggestions in the same visual language as the initial film. This is equivalent to teaching a child how to ride a bicycle by suggesting that they turn the steering wheel in the direction they are falling. It is not the exaggerated move that an experienced bicycle rider would make, but it moves the learner toward the correct goal.

Difficulty: Walking involves a very large number of degrees of freedom, and a much larger possible set of input variables. The recording helps restrict the state space somewhat, but does little to curb the search for appropriate trigger variables. Much of the difficulty comes in reigning in all the parameters. Initially, the system will only use simulated two-dimensional walking, as there are fewer variables, and it is easier to perform.

Impact: Learning by demonstration can be much more straightforward and useful than having to develop an algorithm from scratch for an individual robot. In addition, the tool may be useful for physically realistic animation.

Future Work: Waiting in the Leg Lab is a physical three-dimensional 12 degree of freedom robot named M2. It would be delightful if we can get this demonstration learning system working on the physical robot. Unfortunately, the learning process involves a lot of falling down, so some better safety rigging would have to be developed before trying this. Trying this in simulation first makes the most sense.

Research Support: This research was support by the Tactical Mobile Robotics program at DARPA.

References:

[1] J. Pratt and P. Dilworth and G. Pratt Virtual Model Control of a Biped Walking Robot. *IEEE International Conference on Robotics and Automation*, April, 1997

[2] K. Kirai and M.Hirose The development of Honda Humanoid Robot *IEEE International Conference on Robotics and Automation*, May 1998.

[3] A. Kun and W. T. Miller Adaptive Dynamic Balance of a Biped Robot using Neural Networks *IEEE ISIC/CIRA/ISAS Joint Conference*, Sep 1998.

[4] S. Schaal Learning from Demonstration. *Advances in Neural Network Information Processing Systems*, MIT Press, 1997.

[5] S. Schaal Memory-Based Robot Learning. *IEEE International Conference on Robotics and Automation*, 3, 1994.