Learning to Play Mini-Golf from Human Demonstration using Autonomous Dynamical Systems

S.M. Khansari-Zadeh MOHAMMAD.KHANSARI@EPFL.CH
Aude Billard AUDE.BILLARD@EPFL.CH
LASA Laboratory, Ecole Polytechnique Federale de Lausanne (EPFL), Lausanne, Switzerland

Abstract

We present a new formulation for autonomous (i.e. time-independent) Dynamical Systems (DS) to perform discrete robot motions with *non-zero velocity* at a given target. The proposed model ensures the convergence of all trajectories to the target, and is inherently robust to perturbations. We evaluated the performance of our proposed method to control right-handed swings in mini-golf.

1. Introduction

Dynamical systems provide a powerful tool for robust control of point to point robot motions from a small set of demonstrations. They ensure high precision in reaching a desired target, yet can be easily modulated to generate new motions in areas not seen before. Previous works on DS highlighted the successful learning of discrete (i.e. point-to-point) robot motions either through time-dependent or time-independent DS (for example see (Pastor et al., 2009; Khansari-Zadeh & Billard, 2011; 2010)). While these works address the fundamental concern when learning DS, i.e. stability, they can only be used for generating motions with zero velocity at the target.

The paper presents a formulation for discrete robot movements that models motions with both zero and non-zero velocities. Specifically, we extend the previous formulation of our non-linear *time-independent* DS (Khansari-Zadeh & Billard, 2011) to model robot motions with *a desired velocity* at the target. This extension allows to learn a considerably wider set of motions ranging from pick-and-place movements to agile robot tasks that require reaching/hitting a given target with a specific speed and direction. The most related work to the proposed method was done by (Kober et al., 2010), where they modify the Dynamic Movement Primitive (DMP) formulation (Pastor et al., 2009) to generate hitting motions. Here, we follow a different direction by attempting to estimate a general form of multi-dimensional autonomous DS, and solving the problem from an alternative approach.

Appearing in *Proceedings of the 28th International Conference on Machine Learning*, Bellevue, WA, USA, 2011.

2. Hitting Motion

Consider a state variable $\xi \in \mathbb{R}^d$ that defines the state of a robotic system. Its evolution in time is governed by an autonomous (time-independent) DS according to:

$$\dot{\boldsymbol{\xi}} = \boldsymbol{F}(\boldsymbol{\xi}; \boldsymbol{\theta}) \quad (1)$$

where $\boldsymbol{\theta}$ is the set of parameters of defining $\boldsymbol{F}$. In our formulation, we decompose Eq. 1 into two terms, and formalize it as a multiplication of a target field $\boldsymbol{E}(\boldsymbol{\xi}; \boldsymbol{\theta}_E)$ and a strength factor $v(\boldsymbol{\xi}; \boldsymbol{\theta}_v)$:

$$\dot{\boldsymbol{\xi}} = \boldsymbol{F}(\boldsymbol{\xi}; \boldsymbol{\theta}) = v(\boldsymbol{\xi}; \boldsymbol{\theta}_v)\boldsymbol{E}(\boldsymbol{\xi}; \boldsymbol{\theta}_E) \quad (2)$$

The structure of Eq. 2 is analogous to many physical principles where the motion of a particle in space can be defined with the value of field (e.g. gravity, electrical field, etc.) times a scalar (e.g. mass, electric charge, etc.). The former is a property that describes the space that surrounds a particle, and the latter defines the characteristics of the particle. Similarly, in Eq. 2, the target field describes the form of a motion and the strength factor determines its intensity. The DS parameters $\boldsymbol{\theta}_v$ and $\boldsymbol{\theta}_E$ can be learned based on the user demonstrations. However, these parameters should be estimated such that they 1) ensure the accomplishment of the task starting from any point in space (i.e. global convergence), and 2) generate robot motions that follow the human demonstrations accurately.

To achieve our goal of having a robot motion that produces trajectories that always pass through the target point but with a *non-zero* velocity, we propose to model the target field as follows:

$$\boldsymbol{E}(\boldsymbol{\xi};\boldsymbol{\theta}_E) = \frac{\hat{\boldsymbol{f}}(\boldsymbol{\xi};\boldsymbol{\theta}_E)}{\|\hat{\boldsymbol{f}}(\boldsymbol{\xi};\boldsymbol{\theta}_E)\|} \quad (3)$$

where $\hat{\boldsymbol{f}}(\boldsymbol{\xi};\boldsymbol{\theta}_E)$ is a globally asymptotically stable DS that is learned by Stable Estimator of Dynamical Systems (SEDS) (Khansari-Zadeh & Billard, 2011; 2010). Eq. 3 corresponds to a field with a constant intensity (i.e. $\|\boldsymbol{E}(\boldsymbol{\xi};\boldsymbol{\theta})\| = 1$), and is defined for any point in space except the target. To avoid the singularity at the target, we consider the target field $\boldsymbol{E}(\boldsymbol{\xi};\boldsymbol{\theta}_E) = \boldsymbol{u}_d$, where $\boldsymbol{u}_d$ is a unit vector corresponding to the desired hitting direction. The streamlines of the target field $\boldsymbol{E}$ exactly coincide with those from the SEDS model $\hat{\boldsymbol{f}}$; however in contrast, the value of $\boldsymbol{E}$ is always constant and equal to 1. Hence it ensures all trajectories will path through the target with a non-zero velocity. While the target field can be used to indicate the correct direction of movement, the strength factor $v(\boldsymbol{\xi};\boldsymbol{\theta}_v)$ defines the speed of movement along that direction. Thus, the combination of these two term will result in generating trajectories with similar velocity profiles as the demonstrations.

Given a set of demonstrations, the parameters of the target field can be learned with SEDS, and an estimate of the strength factor can be learned using various existing regression techniques , e.g. Gaussian Process Regression (GPR), or Gaussian Mixture Regression (GMR), etc.

Though the function $\boldsymbol{F}(\boldsymbol{\xi};\boldsymbol{\theta})$ in Eq. 2 results in having a non-zero final velocity, it always approaches the target from the same direction as the demonstrations. In order to approach the target with different orientation, one can use a rotation matrix $\boldsymbol{R}(\boldsymbol{u}_c, \boldsymbol{u}_d)$ to map the demonstration hitting direction $\boldsymbol{u}_c$ to a desired hitting direction $\boldsymbol{u}_d$. In this work, we assume the desired hitting direction is given by the user or a higher level planar. The ability to hit the ball with different magnitude can also be obtained by scaling the strength value.

3. Robot Experiments

We evaluated the performance of the presented method on 6 degrees of freedom industrial robot Katana-T arm for controlling right-handed swings in mini-golf. Playing mini-golf is a non-trivial task: Even a small inaccuracy in hitting direction may result in missing the hole. Hitting speed should also be controlled, otherwise the ball may fail to reach the hole or pass over it. Additionally, different initial gestures and the varied positions of both the ball and hole can make the ability to perform a successful putt more challenging.

For this experiment, we collected a set of demonstrations by kinesthetically moving the robot arm to putt the ball into the hole. For all demonstrations, the relative position of the ball and the hole was fixed, and the user only shows the robot different ways of hitting the ball starting from different initial positions (see Figure 1(b)). The reproductions generated from the final optimized model are illustrated in Figure 1(c). Figure 1(d)-(f) represents the velocity profile of the reproductions versus demonstrations along the axes x, y, and z respectively. In these graphs the thick dashed lines correspond to the position of the ball. The velocity at each point was computed by multiplying the strength factor $v(\boldsymbol{\xi})$ with the target field $E(\boldsymbol{\xi})$. Figure 1(g) shows the sequence of the motion for one of the reproductions. For each reproduction, after hitting the ball, the dynamics were switched to a stable dynamics guiding the arm into a resting position. For this experiment, we considered a simple resting motion where the velocity of the arm end-effector gradually decreases along the direction of the motion until it stops at the resting point.

By considering the frame of reference on the ball and given a correct hitting direction, the above model can also generalize the task to different positions of the hole. Furthermore, Similar to the SEDS model, the proposed model is inherently robust to external perturbations. The recordings of the above robot experiment, its generalization ability to different ball and hole positions, and the evaluation of its robustness to perturbations are provided in:

`http://lasa.epfl.ch/khansari/ICML11_miniGolf.mp4`

4. CONCLUSIONS AND FUTURE WORK

In this work, we presented a novel approach to generating robot motions with a desired velocity at the target. The proposed model retains all advantages of a SEDS model (Khansari-Zadeh & Billard, 2011): i.e. it is globally convergent to the target and inherently robust to external perturbations. The experiment results demonstrated the ability of the proposed model to successfully generalize the task from different initial positions and for various positions of the ball. The robustness of the model was also verified in the presence of perturbations. As for future work, we are currently working on using the proposed method in challenging sports tasks such as playing mini-golf on rough terrains with several obstacles.

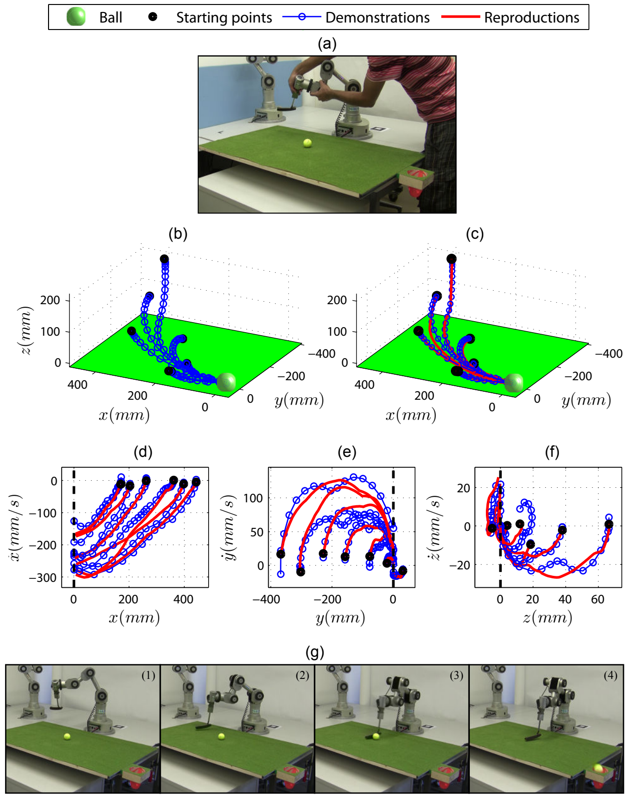

Figure 1. **(a)** Kinesthetic demonstration of putting motion to the 6-DoF Katana-T robot. **(b)** Illustration of the collected successful demonstrations. **(c)** Reproductions of the motion from the model learned with the extended version of SEDS. **(d)-(f)** Evaluation of the model's accuracy in estimating the desired velocity profile. The thick dashed lines locate the position of the ball. **(g)** Illustration of one of the generated motions sequences.

Acknowledgments

This work was supported by the European Commission through the EU Project AMARSI (FP7-ICT-248311).

References

Khansari-Zadeh, S.-M. and Billard, Aude. Imitation learning of globally stable non-linear point-to-point robot motions using nonlinear programming. In *Proceeding of the 2010 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 2676–2683, 2010.

Khansari-Zadeh, S.M. and Billard, A. Learning stable non-linear dynamical systems with Gaussian Mixture Models. *IEEE Transactions on Robotics*, 27, 2011.

Kober, Jens, Mulling, Katharina, Kromer, Oliver, Lampert, Christoph H., Scholkopf, Bernhard, and Peters, Jan. Movement templates for learning of hitting and batting. In *Proceeding of the International Conference on Robotics and Automation (ICRA)*, 2010.

Pastor, P., Hoffmann, H., Asfour, T., and Schaal, S. Learning and generalization of motor skills by learning from demonstration. In *International conference on robotics and automation (ICRA)*, pp. 1293–1298, 2009.