

**Алгоритм обучения систем с дискретным управлением**

**Научный руководитель – Алисейчик Павел Александрович**

**Голиков Кирилл Анатольевич**

*Аспирант*

Московский государственный университет имени М.В.Ломоносова,  
Механико-математический факультет, Кафедра математической теории  
интеллектуальных систем, Москва, Россия

*E-mail: golikov.ka@mail.ru*

Робот однажды поставленный на рабочее место начинает непрерывно работать, выполняя разные позиционирующие движения. Со временем в движущем механизме и окружающей среде происходят изменения, которые отражаются на качестве операций робота. Алгоритм предназначен компенсировать эти изменения, сохраняя точность позиционирования на требуемом уровне.

Причина, по которой робот подвержен указанной переменчивости в том, что он был составлен из некачественных деталей, непредназначенных для выполнения подобных манипуляций с требуемой точностью. Робот состоит из твёрдых конструкций и кинематических пар, к которым подключены отдельные пневмоприводы. Конструкции не являются абсолютно жёсткими, в кинематических связях наблюдаются прогибы и проскальзывания.

Чтобы моделировать данные особенности рассматривается следующая модель.

Модель мускула робота.

**Функционирование мускула** (привода) робота происходит согласно аддитивно-мультипликативной форме случайного процесса в виде композиции:

$$\mathbf{F}(t) = \max(d(t) \mathbf{w}(t) + \mathbf{f}(t), \mathbf{r}(t)) + \mathbf{s}(t) \quad (1)$$

**1) Общие ограничения  $\mathbf{r}(t)$**  на движение мускула робота устанавливаются детерминистической функцией (2) мускула (при её смене ограничения меняются):

- На ускорение - в виде линейной функции  $\mathbf{a}(t) = \mathbf{c}_1 t + \mathbf{c}_2$
- На скорость - в виде максимальной постоянной величины:  $v_{max} = \mathbf{c}_3$

Ограничения мускула  $\mathbf{r}(t)$  полностью определяются коэффициентами  $\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3$ . Эти ограничения, как видно из (1), влияют на все остальные составляющие (кроме выбросов) функции мускула  $\mathbf{F}(t)$ . В очередной такт общая скорость и ускорение мускула с учётом всех составляющих (кроме выброса) не может превышать этих ограничений.

**2) Детерминистическая составляющая  $d(t)$**  - непрерывная (нестрого) монотонно возрастающая функция. Состоит из 2х компонент: перемещение по *усилию* и перемещение по *инерции*.

$$d(t) = \mathbf{Force}(t) + \mathbf{Inertia}(t) = \mathbf{u}(t)( \mathbf{q}(t-1) + \Delta \mathbf{q}(t) ) + \mathbf{u}(t)( \mathbf{c}_4(m_q) \mathbf{q}(t-1) ) \quad (2)$$

где  $\mathbf{c}_4 < 1$  - коэффициент инерционного замедления движения, зависит от  $m_q$  - массы навешенных на мускул элементов (звена, связи, иных звеньев, груза и пр).  $\mathbf{q}(t)$  перемещения мускула произвольная возрастающая (может быть экспонента, линейная, квадратичная и т.д.).

Функция  $d(t)$  жёстко сохраняет свои свойства до очередной смены. *Резкая замена  $d(t)$*  происходит по причинам доступным наблюдению пользователя снаружи робота, примеры этих причин:

- *Пневмопривод*: изменение рабочего тела или рабочего давления, замена привода и пр.

- *Мобильный робот*: изменение поверхности движения, замена катков или гусениц и пр.
- *Манипулятор*: навешивание груза на конструкцию, изменение длины звеньев и пр.

3) **Высокочастотная случайная составляющая  $f(t)$**  - непрерывные затухающие колебания при начале работы мускула с обязательной нормализацией через некоторое число тактов, происходят по ненаблюдаемым причинам в связи с переходными процессами в пневматической работе привода. Колебательность обеспечивается представлением:

$$f(t) = K (1 - A e^{\alpha t} \sin(\omega t + \phi)) \quad (3)$$

где  $c_5(t)$  - передаточный коэффициент,  $c_6(t)$  - коэффициент затухания,  $c_7(t)$  - коэффициент времени. Коэффициенты определяют эту составляющую, могут меняться со временем.

Особенности внутренней работы пневматики, трение поршня о цилиндр, неидеальная передаточные механизмы сил от приводов к связям, особенности кинематических пар и др.

4) **Низкочастотная случайная составляющая  $w(t)$**  - гладкая плавно и медленно меняющаяся в случайно направлении функция долговременных особенностей функционирования мускула, вызванная ненаблюдаемыми систематическими причинами как внешними, так и внутренними. Настолько плавная, что в рамках одного движения (сотни тактов) работа множитель  $w(t) = \text{const}$  можно считать постоянной.  $w(t)$  задаётся произвольным распределением случайной величины (пусть нормальный), характеризуется мат.ожиданием  $m_w = c_8(t)$  и дисперсией  $D_w = c_9(t)$ .

Примеры причин: изменение окружающей температуры и атмосферного давления, износ подшипников сочленений, притёртости в соприкосновении поршня к цилиндру пневмопривода и др.

5) **Выбросы  $s(t)$**  - не является результатом работы привода, имеет строго внешние причины (порыв ветра, удар о препятствие и т.д.), организует разрыв первого рода функции  $F(t)$  в случайный момент времени, на величину  $0 \leq s(t) \leq S_{max}$ .

Невозможно предугадать и предотвратить, необходимо отфильтровать и по возможности нивелировать в процессе осуществления движения, при котором выброс произошёл.

Таким образом, алгоритм обучения будет фильтровать выбросы, подбирать доступные для изучения нестохастические параметры (коэффициенты  $c_1, c_2, c_3, c_4$ ) и мониторить стохастические параметры (коэффициенты  $c_5(t), c_6(t), c_7(t), c_8(t)$ ), создавая несколько моделей сценариев работы мускула, переключаясь между этими сценариями на лету, во время выполнения роботом очередного движения, для обеспечения необходимой точности позиционирования.

Модель работа в данной работе описывается неоднородным случайным Марковским Процессом Принятия Решений (МППР) с дискретными состояниями и дискретным временем. Неоднородный случайный процесс характеризуется зависящими от момента времени

вероятностями перехода между состояниями.

Чтобы задать такой МППР нужно выделить понятия состояния и действия абстрактного робота, при чём, напомним, что в марковском процессе каждое текущее состояние полностью определяется предыдущим состоянием и выбранным действием.

Состояние робота в каждый такт дискретного времени может быть описано изменениями (координат  $(x,y)$  точек на плоскости) положений сочленений (одна из точек сочленений - «особая точка», её правильное положение сигнализирует о выполнении задачи).

У каждого сочленения робота есть набор мускулов, которые перемещают его и, возможно, другие сочленения, на плоскости. Чтобы любое действие модели робота было обратимо, к сочленению приписываем ровно 2 мускула, которые перемещают позиции в прямом и обратном направлениях, имеющие одинаковые усилия.

Действие робота в заданный момент времени можно записать, как целочисленное значение, меньшее  $2^8$ , т.к. число мускулов на одно сочленение 2, а число сочленений не  $\leq 8$ .

Использованный алгоритм для обучения системы есть вариант *метода спуска* (hill climbing) с выбором первого найденного среди соседних решения, улучшающего результат, и непосредственного перехода в него. В нашем случае, *функция приспособленности*, определяющая качество решения относительно остальных решений, в задаче позиционирования «особой точки» робота в фиксированную цель представляется хорошей вогнутой поверхностью с единственным минимумом.

Алгоритм разрабатывается, чтобы исключить зависимость обхода от типа голономности робота.

### Источники и литература

- 1) Яблонский С.В. Введение в дискретную математику. — М.:Высшая школа, 2006.
- 2) Гасанов Э.Э., Кудрявцев В.Б. Интеллектуальные системы. Теория хранения и поиска информации — 2-е изд., испр. и доп. — М.:Юрайт, 2017.
- 3) Саттон Р.С., Барто Э.Г. Обучение с подкреплением — 2-е изд. — М.:БИНОМ.Лаборатория знаний, 2014.