

УДК 681.513.7

Г. А. ПРОКОПОВИЧ

АДАПТИВНАЯ НЕЙРОСЕТЕВАЯ СИСТЕМА УПРАВЛЕНИЯ АВТОНОМНЫМ МОБИЛЬНЫМ РОБОТОМ НА ОСНОВЕ МЕТОДА ОБУЧЕНИЯ С УЧИТЕЛЕМ В *ONLINE* РЕЖИМЕ

Объединенный институт проблем информатики НАН Беларуси

(Поступила в редакцию 06.03.2015)

Ведение. Одно из лидирующих направлений в области искусственного интеллекта связано с моделированием искусственных нейронных сетей (ИНС), способных решать широкий круг задач распознавания образов, идентификации, прогнозирования, оптимизации и управления сложными объектами [1, 2]. ИНС строятся по принципу организации и функционирования их биологических аналогов [3].

Преимущества ИНС перед традиционными методами искусственного интеллекта (продукционные системы, таблицы принятия решений, генетические алгоритмы, нечеткая логика и т. д.) проявляются при решении таких задач, когда трудно выделить правила функционирования систем, обладающих большим числом степеней свободы, зато существует возможность обучения на множестве примеров [4].

Тем не менее ИНС имеют и свои ограничения, связанные с отсутствием универсальной архитектуры и единых подходов для использования в системах управления, а также с проблемами дообучения новой информации. На данный момент в известных системах управления техническими устройствами ИНС используются в виде отдельных блоков и в основном только для реализации функций классификации или интерполяции данных, что значительно сужает их потенциал.

1. Недостатки традиционных способов обучения ИНС. Возможность обучения – одно из главных преимуществ ИНС перед традиционными алгоритмами машинного обучения, которое основано на свойствах биологической памяти [4–6]. В процессе обучения ИНС способны выявлять сложные зависимости между входными и выходными данными, а также выполнять *обобщение* [6], под которым понимается способность ИНС на основе обучающей выборки генерировать правильный ответ при анализе входных данных, которые не встречались в процессе обучения.

Одним из наиболее часто используемых способов обучения ИНС является обучение с учителем (рис. 1), когда выходное пространство решений нейронной сети известно и заключается в последовательном предъявлении на вход ИНС обучающей выборки с указанием соответствующих

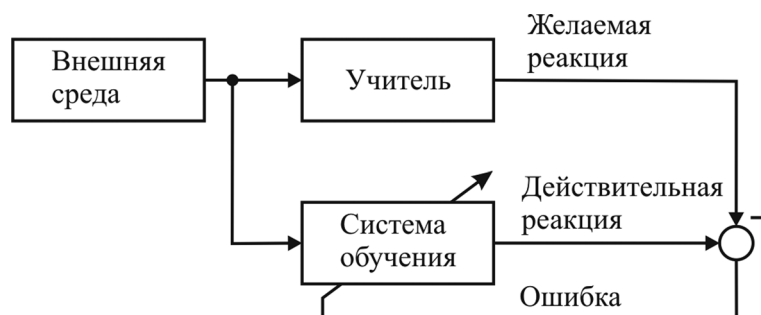


Рис. 1. Схема процесса обучения ИНС с учителем

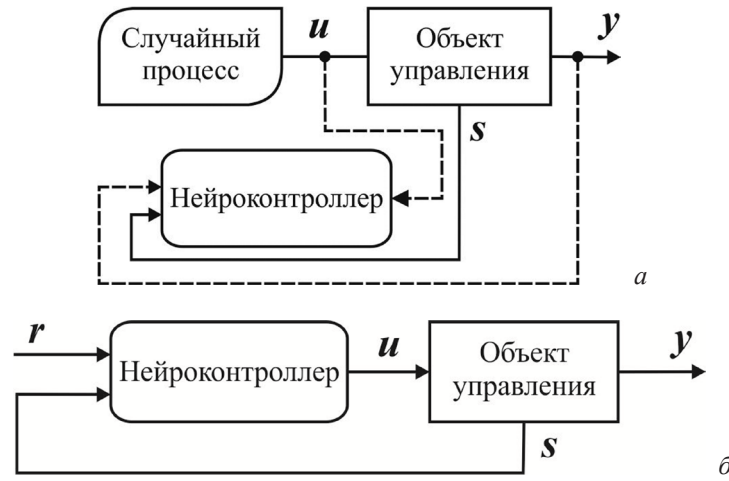


Рис. 2. Схема обобщенного инверсного нейроуправления:
 а – режим обучения инверсного нейроконтроллера; б – режим управления объектом

значений на ее выходах – схема прямого контролируемого обучения [6]. Присутствующий в схеме «учитель» (или эксперт) должен обладать знаниями о внешней среде, представленными в виде набора эталонных образцов вход-выход. Пара таких эталонов называется обучающей выборкой, которая включает в себя, соответственно, входной $\mathbf{x}_i^*$ и выходной $\mathbf{y}_i^*$ векторы. Множество пар эталонов $\{\mathbf{x}_1^*, \mathbf{x}_2^*, \dots, \mathbf{x}_k^*\}$ и $\{\mathbf{y}_1^*, \mathbf{y}_2^*, \dots, \mathbf{y}_k^*\}$ называется обучающим набором $U = \{\mathbf{x}_i^*, \mathbf{y}_i^*\}_{i=1, \overline{k}}$.

Как правило, изначально сама ИНС никакими знаниями не обладает. В процессе обучения эксперт и ИНС подвергаются воздействию внешней среды, т. е. на входы последней поступает тренировочный сигнал $\mathbf{x}_i$, совпадающий с одним из входных эталонов $\mathbf{x}_i^*$. Эксперт сообщает сети, какова должна быть правильная (желаемая) реакция на поступившее воздействие, подавая соответствующий выходной эталон $\mathbf{y}_i^*$. На основании величины ошибки между реальным и желаемым выходами сети по определенному правилу проводится настройка ее синаптических связей [6]. Повторяя этот процесс итеративно, можно настроить ИНС так, что она будет эмулировать знания эксперта о внешней среде.

На рис. 2, а приведена схема инверсного нейроконтроллера, который в процессе обучения с учителем способен в некоторых пределах сформировать модель поведения исследуемого объекта [7]. Инверсный нейроконтроллер представляет собой ИНС, обученную в режиме *offline* имитировать обратную динамику объекта управления на основе выявления фазовых траекторий поведения динамического объекта.

Для получения таких траекторий на объект управления в качестве управляющего сигнала подается некоторый случайный процесс. Значения управляющих сигналов $\mathbf{u}^{(N)} = (u_1, u_2, \dots, u_N)^T$ и ответных реакций объекта $\mathbf{y}^{(N)} = (y_1, y_2, \dots, y_N)^T$ и $\mathbf{s}^{(M)} = (s_1, s_2, \dots, s_M)^T$ протоколируются, и на этой основе формируется обучающая выборка для нейроконтроллера:

$$U = \{P_i, T_i\}_{i=1, \overline{k}}, \quad \begin{cases} P_i^{(N+M)} = \begin{bmatrix} \mathbf{y}^{(N)}[n] \\ \text{---} \\ \mathbf{s}^{(M)}[n-1] \end{bmatrix}, \\ T_i^{(N)} = \mathbf{u}^{(N)}[n-1], \end{cases} \quad (1)$$

где $P_i^{(N+M)}$ – вектор-столбец входных сигналов; $T_i^{(N)}$ – вектор-столбец выходных сигналов; n – шаг дискретного времени; k – число обучающих выборок; $\mathbf{s}^{(M)}$ – вектор, описывающий состояния объекта, причем, как правило, для оценки текущего состояния объекта управления используют последовательность из его прошлых реакций на входные воздействия

$$\mathbf{s}^{(M)}[n] = \left(\mathbf{y}[n-1]^T, \mathbf{y}[n-2]^T, \dots, \mathbf{y}\left[n - \frac{M}{N}\right]^T \right)^T, \quad (2)$$

либо мгновенный срез его фазовых траекторий

$$\mathbf{s}^{(M)}[n] = \left(\mathbf{y}[n]^T, \mathbf{y}'[n]^T, \mathbf{y}''[n]^T, \dots, \mathbf{y}^{(p)}[n]^T \right)^T, \quad (3)$$

где $M \geq N$, причем N кратно M ; p является номером порядка производной, которая равняется $p = \frac{M}{N} - 1$.

В ходе обучения ИНС должна определить и запомнить зависимость значений управляющего сигнала $\mathbf{u}[n-1]$ от последующего значения реакции объекта управления $\mathbf{y}[n]$, находящегося перед этим в состоянии $\mathbf{s}[n-1]$. Во время управления объектом (рис. 2, б) инверсный нейроконтроллер подключается таким образом, чтобы на вход $\mathbf{x}[n]$ подавались значения установки $\mathbf{r}[n]$ и состояния объекта управления $\mathbf{s}[n-1]$, поступающие по каналу обратной связи:

$$\mathbf{x}[n] = \begin{bmatrix} \mathbf{r}[n] \\ \text{-----} \\ \mathbf{s}[n-1] \end{bmatrix}. \quad (4)$$

Предполагается, что сформированная при обучении инверсная модель объекта управления является адекватной реальному объекту. Тогда сигнал управления $\mathbf{y}[n+1]$, получаемый с выхода ИНС, обеспечивает переход объекта в стационарное состояние, которое было задано экспертом в процессе обучения.

Однако указанный способ обучения чаще всего используется в *offline* режиме, так как в процессе дообучения большинства ИНС новым образом возникает проблема, которая является одной из самых сложных и трудно решаемых при построении интеллектуальных систем, моделирующих распознавание образов и восприятие окружающего мира [5, 6]. Она заключается в определении баланса между процессами синаптической *пластичности* нейронов для запоминания новых и *стабильного* хранения уже существующих в памяти образов.

В некоторых случаях процесс «затирания» старой информации является несущественным. Если набор обучающих векторов фиксирован, то, например, процесс обучения многослойного персептрона по методу обратного распространения ошибки будет заключаться в многократном циклическом предъявлении обучающей выборки. В итоге персептрон способен запомнить весь набор обучающих выборок. Однако при обучении персептрона новому образу произойдет модификация синаптических связей с неконтролируемым разрушением структуры памяти о предыдущих образах. Таким образом, персептронная модель ИНС не способна к запоминанию новой информации *без полного переобучения сети* [6].

Для решения указанной проблемы в работе [8] автором был предложен адаптивный нейросетевой классификатор АРТ-ДАП с функцией дообучения, который позволяет в отличие от известных аналогов, основанных на адаптивной резонансной теории С. Гроссберга, производить ассоциативный поиск хранящихся в памяти эталонов. Следующим немаловажным свойством разработанного классификатора является возможность реализации процесса обучения с учителем, что позволяет эксперту объединять в одном классе различные эталоны [9]. Данные свойства достигаются благодаря реализации долговременной памяти классификатора на основе архитектуры гетероассоциативной ИНС – двунаправленно ассоциативной памяти (ДАП), а также оригинальному алгоритму выявления новых образов.

2. Процесс самообучения нейросетевых систем управления на основе реализации механизма условных рефлексов. Решение задачи движения автономных мобильных роботов (АМР) в априори неизвестной местности с помощью классических подходов теории управления, как правило, является проблематичным. Это связано с тем, что не всегда можно подобрать точную аналитически заданную форму функциональных зависимостей входных и выходных сигналов системы «датчики–двигатели». В качестве альтернативы существуют подходы проектирования

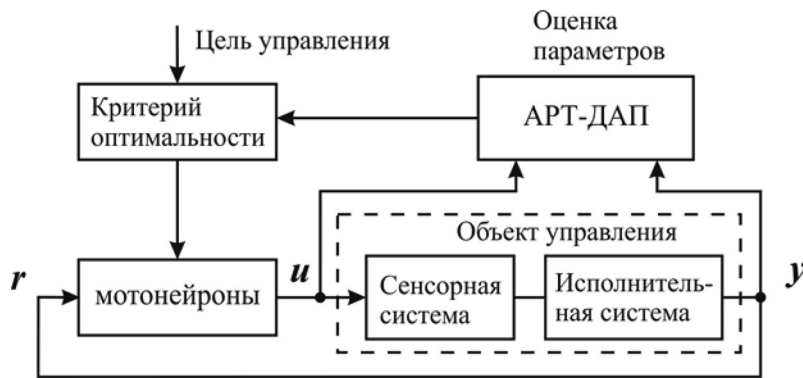


Рис. 3. Функциональная схема предложенного адаптивного нейроконтроллера

управляющих систем на основе самообучения, которые, обучаясь на опыте своего взаимодействия с внешней средой, способны самостоятельно формировать свое поведение [2–4]. Обычно подобные системы демонстрируют значительно более высокие возможности обучения и адаптации по сравнению с классическими инженерными подходами. Причем в большинстве случаев целевые функции априори содержатся в синаптических связях самих ИНС, т. е. $r[n+1] = y[n]$.

Для приспособления к постоянным условиям функционирования, при которых параметры среды изменяются в незначительных пределах, наиболее экономным средством адаптации как для естественных, так и для искусственных систем управления являются механизмы условных и безусловных рефлексов, позволяющие ограничиться весьма простым набором правил [3].

Таким образом, одним из самых простых, но не менее ответственных среди функциональных блоков проектируемых систем управления АМР должен быть реактивный контур управления, предназначенный для реализации аппарата безусловно-условных рефлексов [10]. Реактивный контур управления должен обладать способностью к обучению и запоминанию на определенное время определенных реакций исполнительных систем АМР при поступлении соответствующих стимулов. Для обеспечения перечисленных функциональных свойств его блок памяти должен выполнять следующие действия: восстанавливать из памяти и классифицировать входные образы на основе уже имеющихся эталонов, а также добавлять новые [9].

В настоящей работе предлагается архитектура адаптивного нейроконтроллера, состоящего из двух гетероассоциативных ИНС (рис. 3). Первая ИНС реализована в виде блока моторнейронов, в котором хранится алгоритм управления исполнительными системами робота. Задачей второй ИНС, в качестве которой используется разработанный ранее классификатор АРТ-ДАП, является выполнение функции контроля процесса дообучения блока моторнейронов. В результате нейроконтроллер способен не только реализовывать выполнение наперед заданной системы правил, но и приспособливаться к новым внешним условиям.

На рис. 4 приводится структурная схема контура реактивного управления АМР, основанная на принципах работы предложенного адаптивного нейроконтроллера, которая работает следующим

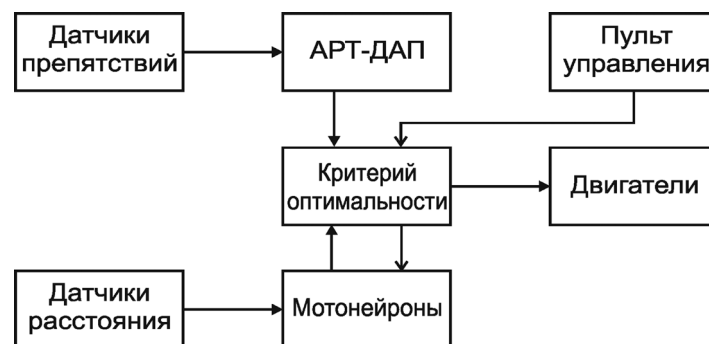


Рис. 4. Структурная схема реактивного контура управления для АМР, состоящего из систем безусловных и условных рефлексов

образом. В блок АРТ-ДАП, который отвечает за реализацию безусловных рефлексов, до начала функционирования АМР экспертом должны быть записаны ассоциативные пары «стимул-реакция», представляющие собой набор критических значений ответственных параметров $\{x_i^*\}_{i=\overline{1,k}}$ мобильного робота и соответствующий им набор ответных действий $\{y_i^*\}_{i=\overline{1,k}}$ исполнительной системы, которые предназначены для ее перехода обратно в нормальный режим. В блоке мотонейронов должны содержаться ассоциативные пары «стимул-реакция», однако (в отличие от системы безусловных рефлексов) в них хранится набор действий исполнительной системы, функционирующей в штатном режиме. Следует отметить, что размер входного вектора для блока мотонейронов значительно превышает размер входного вектора для блока АРТ-ДАП. Это связано с различными типами используемых датчиков в каждом из описываемых блоков. Например, на вход блока АРТ-ДАП могут поступать цифровые сигналы датчиков препятствий (ближняя зона), а на вход блока мотонейронов – сигналы сканирующего дальномера (рабочая область), причем сигналы сканирующего дальномера представляют собой набор большого числа аналоговых значений.

Принцип работы блока критерия оптимальности заключается в следующем: если управляющие значения $u_{\text{мото}}^{(N)}$, полученные от блока мотонейронов, не противоречат управляющим значениям $u_{\text{АРТ-ДАП}}^{(N)}$ блока безопасного движения АРТ-ДАП, то блок критерия оптимальности соединит выходы блока мотонейронов с входами блока драйвера двигателей.

В противном случае блок критерия оптимальности принудительно создаст в долговременной памяти ИНС мотонейронов $(k+1)$ -ю ассоциативную пару, выходной обучающий вектор $T_{k+1}^{(N)}$ которой будет представлять собой управляющий вектор от блока АРТ-ДАП, а входной – обучающий вектор $P_{k+1}^{(N+M)}$, содержащий текущие данные от сканирующего дальномера и состояние объекта управления q шагов модельного времени n назад. При повторении данной ситуации блок мотонейронов выдаст такие управляющие сигналы, которые не приведут к появлению внештатного режима работы АМР (в описываемом случае – столкновение с препятствием). Другими словами, обучающая выборка $U_{\text{мото}} = \{P_i, T_i\}_{i=\overline{1,k+1}}$ по формуле (1) для ИНС мотонейронов предложенного адаптивного нейроконтроллера будет иметь следующий вид:

$$\begin{cases} u_{\text{мото}}^{(N)}[n] \neq u_{\text{АРТ-ДАП}}^{(N)}[n], \\ P_{k+1}^{(N+M)} = \begin{bmatrix} y^{(N)}[n-q] \\ \text{-----} \\ s^{(M)}[n-q-1] \end{bmatrix}, \\ T_{k+1}^{(N)} = u_{\text{АРТ-ДАП}}^{(N)}[n-1], \\ q = \frac{L}{V \cdot \Delta n}, \end{cases} \quad (5)$$

где L – разница между радиусами действия сенсоров рабочей и ближней зон АМР, значения которых поступают на входы блоков мотонейронов и АРТ-ДАП соответственно; V – скорость движения АМР; Δn – шаг дискретного модельного времени n , причем q округляется к большему целому числу.

Таким образом, в процессе функционирования АМР в *online* режиме в блок мотонейронов могут добавляться новые ассоциативные пары «стимул-реакция». Данный процесс можно рассматривать как генерацию условных рефлексов, позволяющих АМР адаптироваться к новым условиям функционирования.

Для отладки и тестирования предложенного нейроконтроллера нами была разработана программа имитационного моделирования, способная воспроизвести работу типового автономного мобильного робота, сенсорная система которого состоит из датчиков безопасности и сканирующего дальномера [11]. Показано, что предложенный адаптивный нейроконтроллер, согласно

заранее заданному в блоке АРТ-ДАП алгоритму движения, был способен не только генерировать управляющие сигналы, но и вырабатывать в процессе функционирования в блоке мотонейронов индивидуальные реакции АМР на некоторые внешние стимулы. Это позволило АМР за 4–6 соударений (попадания различных преград в его зону безопасности) научиться двигаться по плоскости, не задевая препятствия, причем опираясь уже на данные от сканирующего дальномера.

Полученные результаты сравнивались с наиболее близкой по назначению нейроноподобной системой автономного адаптивного управления (ААУ), описанной в работе [2] и предназначенной для управления мобильным роботом с аналогичной сенсорной системой. В приведенных в [2] экспериментах системе управления требовалось 10–20 столкновений, чтобы база знаний накопила достаточно примеров и далее робот двигался без столкновений.

Заключение. На основе реализации механизмов безусловных и условных рефлексов реализована нейросетевая система управления автономным мобильным роботом, которая предназначена не только для генерации управляющих сигналов, но и выработки в процессе функционирования робота индивидуальных реакций на некоторые внешние стимулы. При этом объем поступающих из внешней среды сенсорных данных передается в соответствующие управляющие блоки дифференцированно: для активизации безусловных рефлексов требуется значительно меньшее количество информации, чем для условных. Указанные свойства предложенной системы управления позволяют роботу эффективно адаптироваться к изменениям внешней среды.

Благодаря использованию архитектур гетероассоциативных ИНС была реализована схема процесса обучения блока мотонейронов с учителем в *online* режиме. Величина временного сдвига q при генерации обучающей выборки зависит от характеристик используемых сенсоров, скорости движения мобильного робота, а также временного шага модельного времени.

Работа выполнена при частичной поддержке гранта БРФФИ-ГФФИУ № Ф13К-144 «Разработка методов оперативной обработки и передачи информации для эффективного управления мобильными роботами и подвижными системами».

Литература

1. Новые методы управления сложными системами. М., 2004.
2. Жданов А. А. Автономный искусственный интеллект. М., 2009.
3. Гаазе-Рапопорт М. Г. От амебы до робота: модели поведения. М., 1987.
4. Амосов Н. М. Нейрокомпьютеры и интеллектуальные роботы. Киев, 1991.
5. Рассел С. Искусственный интеллект: современный подход. М., 2006.
6. Хайкин С. Нейронные сети: полный курс. М., 2006.
7. Омату С. Нейроуправление и его приложения. М., 2000.
8. Прокопович Г. А. // Информатика. 2009. № 3 (23). С. 68–81.
9. Прокопович Г. А. // Информатика. 2010. № 2 (26). С. 54–65.
10. Прокопович Г. А. // Искусственный интеллект. 2013. № 1. С. 181–190.
11. Прокопович Г. А. // Информатика. 2014. № 4 (44). С. 34–43.

R. A. PRAKAPOVICH

ADAPTIVE NEURAL NETWORK CONTROL SYSTEM OF AN AUTONOMOUS ROBOTIC VEHICLE BASED ON *ONLINE* SUPERVISED LEARNING

Summary

The disadvantages of the classical architectures of artificial neural networks (ANNs) in the problems of intelligent control of an autonomous robotic vehicle are described. Based on updated bi-directional associative ANNs an adaptive neurocontroller has been developed which enables one to find the cause-effect relationships in the “robot–environment” system. The neurocontroller is based on the rule-based system and contains two ANNs that perform two different functions. The first one is implemented as a motoneurons unit that contains the robot motion control algorithm, and the second one is designed to identify in the sensory data sequence new patterns that are added to the first ANN based on the supervised learning scheme.