

$$+2\alpha \cdot \frac{1}{L^3} \cdot \left(\sum_{j_0=1}^{m_0+1} \sum_{j_1=1}^{m_1} \sum_{l_1=1}^{m_1+1} \sum_{l_2=1}^{m_2} G_{l_1 l_2}^{(2)} \cdot (S^{(1,2)})_{l_1 l_2}^{j_0 j_1} \cdot G_{j_0 j_1}^{(1)} \right) = 0.$$

Тогда шаг обучения α определяется в соответствии с выражением:

$$\alpha = L \cdot \frac{D_1 + D_2}{A + 2B + C},$$

А значит

$$\alpha_m = \alpha \cdot \frac{1}{L} = \frac{D_1 + D_2}{A + 2B + C}$$

Теорема доказана.

4. ЗАКЛЮЧЕНИЕ

В данной статье предложена методика обучения двух-слойных гетерогенных сетей прямого распространения на основе метода наискорейшего спуска: послойное обучение, двухпараметрическое обучение и обобщенный метод наискорейшего спуска. Для каждого из этих подходов приведены и доказаны соотношения для изменения синаптических связей и формулы для вычисления адаптивных шагов. Эксперименты, проведенные на тестовых данных, показывают более быструю сходимость метода двухпараметрического обучения, хотя каждый из выше предложенных методов может дать

УДК 681.3

Головко В.А., Дунец А.П., Кирись А.Н., Селезнев П.В.

РАСПОЗНАВАНИЕ РЕГИСТРАЦИОННЫХ НОМЕРОВ АВТОМОБИЛЕЙ С ИСПОЛЬЗОВАНИЕМ НЕЙРОННЫХ СЕТЕЙ

Введение

Автоматическое распознавание регистрационных номеров автомобилей в реальном режиме времени является актуальной задачей, решение которой может использоваться во многих практических приложениях. Это, например, контроль автомобилей на автострадах или при парковке на автостоянках. При этом предполагается, что имеется видеокамера, которая отображает изображение регистрационного номера автомобиля. Изображение является зашумленным. Задачей системы распознавания является корректная идентификация номера автомобиля в реальном масштабе времени и представление его в текстовом формате.

В настоящее время имеется много научных публикаций и коммерческого программного обеспечения, предназначенного для распознавания регистрационных номеров автомобилей [1,2,3]. Основная проблема здесь заключается в повышении качества распознавания.

В данной статье описываются и анализируются два нейросетевых метода распознавания автомобильных номеров. Первый метод базируется на преобразовании Фурье исходно-

лучший результат в сравнении с двумя другими, при определенных тестовых данных.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Danuta Rutkowska, Maciej Piliński, Leszek Rutkowski. Sieci neuronowe, algorytmy genetyczne I systemy rozmyte. – Warszawa, Łódź: Wydawnictwo Naukowe PWN, 1999. – 412 s.
2. Hornik K., Stinchcombe M. and White H. Multilayer feedforward networks are universal approximators // Neural Networks, 1989, vol.2 – PP. 359-366.
3. Гладкий И.И., Головко В.А., Махнист Л.П. Обучение нейронных сетей с использованием метода наискорейшего спуска // Вестник Брестского государственного технического университета. – Брест: БГТУ, 2001. № 5: Физика, математика, химия. – С. 56-61.
4. Маньяков Н.В., Махнист Л.П. Матричная алгоритмизация обучения многослойных нейронных сетей с использованием градиентных методов // Вестник Брестского государственного технического университета. – Брест: БГТУ, 2002. - №5 (17): Физика, математика, химия. – С. 60-64.
5. V. Golovko, N. Maniakov, L. Makhnist. Multilayer Neural Networks Training Methodic // Proceedings of the Workshop Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS'2003). – Lviv, Ukraine, September 8-10, 2003. – PP. 185-190.

го изображения номера автомобиля и применении многослойного перцептрона для распознавания. Второй подход основывается на применении в качестве распознавателя самоорганизующейся сети Кохонена. Результаты экспериментов обсуждаются.

1. Структура системы распознавания

Рассмотрим общую структуру системы распознавания регистрационных номеров автомобилей. Она состоит из модуля сегментации, модуля обработки изображений и модуля нейронных сетей (рис.1).

Модуль сегментации осуществляет разделение исходного изображения номера автомобиля на отдельные участки, каждый из которых соответствует определённому знаку (букве или цифре) регистрационного номера (рис.2). Для этого можно использовать градиентный метод, описанный в [4].

Модуль обработки производит преобразование изображения номера автомобиля в форму удобную для подачи на модуль нейронных сетей. В данной работе рассмотрено два варианта реализации данного модуля.

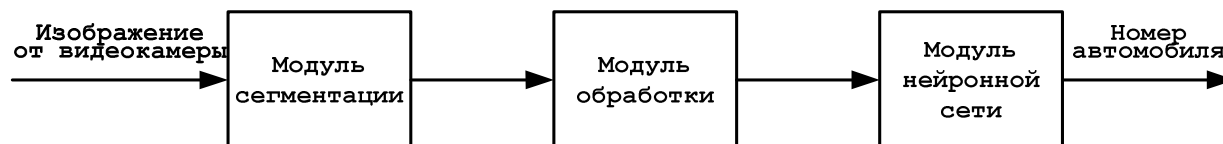


Рис.1. Структурная схема системы распознавания.

Дунец Андрей Петрович, ст. преподаватель каф. интеллектуальных информационных технологий Брестского государственного технического университета.

Кирись А.Н., ст. преподаватель каф. ИИТ Брестского государственного технического университета.

Селезнев Петр Владимирович, ст. преподаватель каф. ИИТ Брестского государственного технического университета.

Беларусь, БГТУ, 224017, г. Брест, ул. Московская, 267.

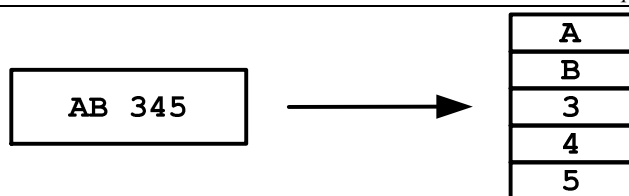


Рис. 2. Сегментация изображения: а - исходное изображение, б - сегментированное.

Первый вариант реализации основан на комбинации простых фильтров. Он осуществляет следующие функции: медианная фильтрация, пороговая обработка, отсеечение и масштабирование. Медианный фильтр используется для очистки изображения от аддитивного и импульсного шумов.

Операция пороговой обработки состоит в преобразовании полутонового изображения в бинарное, которое имеет две градации яркости. В качестве порогового значения в данной работе использовалось среднее арифметическое между математическим ожиданием яркости всего изображения и темной её части. На рис. 3а представлено оригинальное изображение от модуля сегментации, а на рис. 3б – изображение, полученное после пороговой обработки.

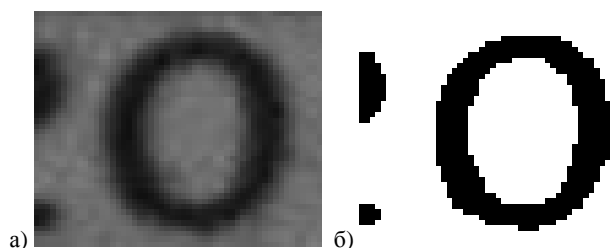


Рис. 3. Исходное изображение (а) и изображение, полученное после пороговой обработки (б).

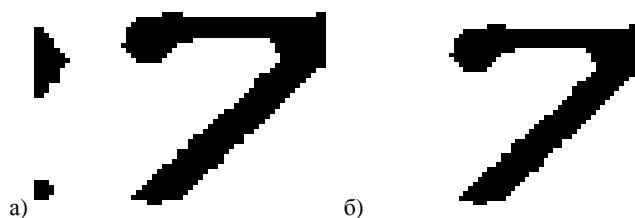


Рис. 4. Операция отсеечения: а – исходное изображение; б – изображение после отсеечения.

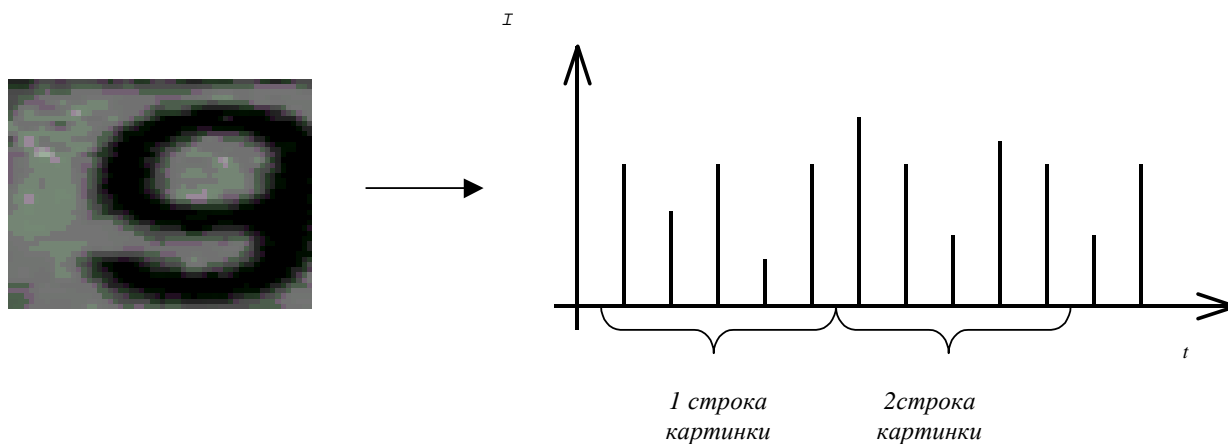


Рис. 6. Преобразование изображения во временной ряд.

Операция отсеечения предназначена для корректировки результатов сегментации. Так после сегментации сегменты могут содержать части знаков других сегментов. Поэтому процедура отсеечения состоит в очистке сегмента от содержащихся в нем остатков знаков других сегментов.

Операция масштабирования состоит в позиционировании знака номера по центру и растяжении его по высоте с целью единообразного представления сегментов знаков (рис.5).

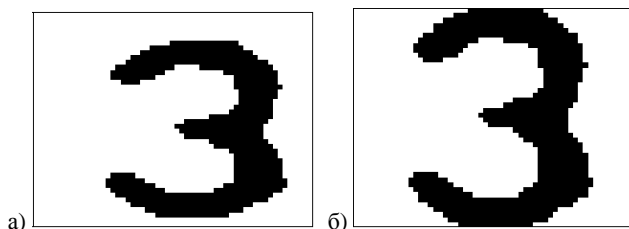


Рис. 5. Исходное (а) и масштабированное (б) изображения.

Второй вариант реализации модуля использует для предварительной фильтрации преобразование Фурье [5]. Изображение одного символа автомобильного номера преобразуется в вектор путем построчной развертки (рис. 6). Значение каждого отсчета – это яркость соответствующего пикселя.

Далее над получившимся вектором выполняется преобразование Фурье

$$\bar{P} = FFT(\bar{I}). \quad (1)$$

В векторе \bar{P} элемент p_0 – постоянная составляющая сигнала. Далее выполняется нормирование вектора P , с целью компенсации влияния разной внешней освещенности автомобильного номера:

$$\bar{P}^* = \frac{\bar{P}}{p_0}. \quad (2)$$

Из получившегося вектора \bar{P}^* на вход нейронной сети подаются только первые 50 элементов.

Модуль нейронных сетей предназначен для непосредственного распознавания регистрационного номера автомобиля. В общем случае он состоит из двух параллельно функционирующих нейронных сетей, одна из которых предназначена для распознавания цифровых, а вторая буквенных знаков. При такой конфигурации на вход первой сети последовательно поступают сегменты цифр, а на вход второй – сегменты букв. Как уже отмечалось, в данной статье для операции распознавания использовалось нейронная сеть Кохонена и многослойный перцептрон. В следующих разделах это описывается более подробно.

2. Нейронная сеть Кохонена

Является самоорганизующейся в процессе работы и осуществляет топологическое упорядочивание входного пространства образов [6]. Она состоит из двух слоев нейронных элементов (рис. 7) и осуществляет разбиение n -мерного входного пространства на различные области, каждой из которой соответствует отдельный нейрон второго слоя.

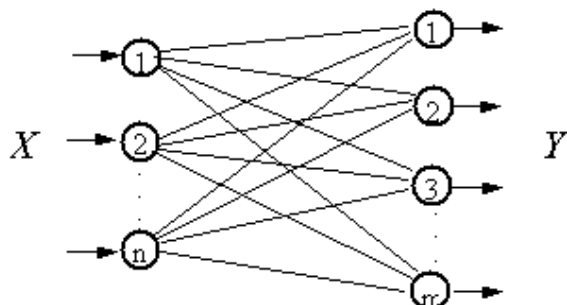


Рис. 7. Нейронная сеть Кохонена.

Для обучения и функционирования самоорганизующейся нейронной сети используется конкурентный метод. В соответствии с ним определяется нейрон победитель с номером k , который соответствует минимальному евклидовому расстоянию между входным и весовым вектором:

$$D_k = \min_j |X - W_j| = \min_j \sqrt{(x_1 - w_{1j})^2 + (x_2 - w_{2j})^2 + \dots + (x_n - w_{nj})^2}, \quad (3)$$

где $X = \{x_1, x_2, \dots, x_n\}$ - входной образ, $W_j = \{w_{1j}, w_{2j}, \dots, w_{nj}\}$ - весовые коэффициенты j -го входного нейрона, n -размерность входного образа.

Тогда весовые коэффициенты k -го нейрона-победителя усиливаются

$$w_{ik}(t+1) = w_{ik}(t) + \gamma(x_i - w_{ik}(t)), \quad (4)$$

где $i=1, n$, а для остальных нейронов весовые коэффициенты не изменяются. На этапе функционирования такой сети выходное значение нейронного элемента победителя принимается равным единице, а выходные значения остальных нейронов равняются нулевому значению. Тогда при поступлении на вход такой нейронной сети неизвестного образа, он идентифицируется в соответствии с мерой близости к эталонным векторам и кодируется на выходе сети номером нейрона.

В данной статье используется контролируемое конкурентное обучение (supervised competitive learning), которое предполагает, что заранее известно соответствие эталонных векторов нейронным элементам. Для такого обучения весовые коэффициенты нейрона победителя усиливаются при корректной классификации и ослабляются в противном случае. Тогда

$$\Delta W_k = \gamma(X - W_k), \quad (5)$$

если X и W_k принадлежат к одному классу и

$$\Delta W_k = -\gamma(X - W_k), \quad (6)$$

если X и W_k принадлежат к различным классам. Весовые коэффициенты остальных нейронов при этом не изменяются.

Для обучения нейронной сети использовалась база данных автомобильных номеров, полученная от видеокамеры. Она содержит около 2000 автомобильных номеров, размером от 640x480 пикселей. На вход сети подавалось бинарное

масштабированное изображение. При этом использовались две нейронные сети: одна обучалась для распознавания цифр, а вторая - букв латинского алфавита.

На первом этапе оценивалось качество распознавания букв латинского алфавита. Распознавание 223 графических образов, содержащих 5-18 изображений букв, дало результат 98.52% (рис.8).

Recognition accuracy:

I:	18 entries	96.80 %
P:	16 entries	97.62 %
M:	0 entries	97.77 %
A:	12 entries	96.02 %
K:	16 entries	99.04 %
O:	17 entries	98.18 %
B:	18 entries	97.91 %
C:	15 entries	96.21 %
D:	13 entries	97.67 %
E:	9 entries	96.58 %
G:	7 entries	99.01 %
Z:	7 entries	97.11 %
N:	9 entries	98.56 %
S:	7 entries	97.62 %
V:	5 entries	99.67 %
T:	8 entries	96.52 %
F:	7 entries	99.43 %
J:	7 entries	97.99 %
U:	8 entries	96.46 %
X:	6 entries	97.18 %
Y:	10 entries	98.05 %

Total accuracy: 223 entries 98.52 %

Рис. 8. Протокол распознавания букв.

На втором этапе производилось распознавание цифр (от 0 до 9). Пример образов для распознавания представлен на рис.9. Цифры представлены 178 образами, от 11 до 21 образа на каждую цифру. Результат распознавания 97.89% (рис. 10).

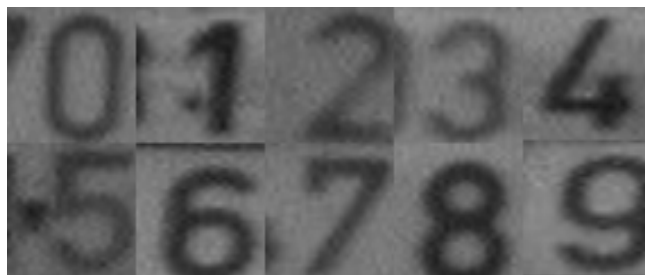


Рис. 9. Образы цифр для распознавания.

Recognition accuracy:

0:	26 entries	99.77 %
1:	21 entries	98.59 %
2:	11 entries	99.03 %
3:	26 entries	96.93 %
4:	12 entries	98.03 %
5:	24 entries	96.09 %
6:	10 entries	98.23 %
7:	22 entries	96.01 %
8:	11 entries	97.11 %
9:	17 entries	99.12 %

Total accuracy: 178 entries 97.89 %

Рис.10. Статистика распознавания цифр.

Таким образом, самоорганизующаяся сеть Кохонена осуществляет достаточно точное распознавание номерных знаков автомобилей.

3. Многослойный перцептрон

Архитектура многослойной нейронной сети состоит из множества слоев нейронных элементов (рис. 11) [6].

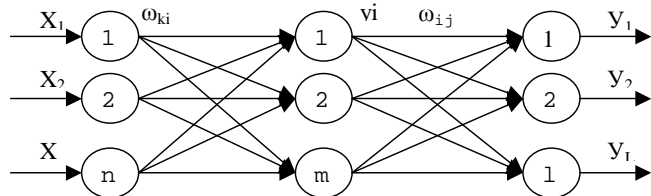


Рис. 11. Трехслойная нейронная сеть.

Входной слой (input layer) нейронных элементов выполняет распределительные функции. Выходной слой (output layer) нейронов служит для обработки информации от предыдущих слоев и выдачи результатов.

Слой нейронных элементов, расположенный между входным и выходным слоем, называется промежуточным или скрытым (hidden layer). Как и выходной слой, скрытый слой является обрабатывающим. Выход каждого нейронного элемента предыдущего слоя нейронной сети соединен синаптическими связями со всеми входами нейронных элементов следующего слоя. Таким образом, топология многослойной нейронной сети является однородной и регулярной.

В качестве функции активации нейронных элементов использовалась сигмоидная функция.

Обозначим слои нейронных элементов от входа к выходу соответственно через k, i, j . Тогда выходное значение l -го нейрона последнего слоя равняется:

$$y_j = F(S_j), \tag{7}$$

$$S_j = \sum_{i=1}^m \omega_{ij} y_i + T_j, \tag{8}$$

где S_j - взвешенная сумма j -го нейрона выходного слоя; y_i - выходное значение j -го нейрона предпоследнего слоя; ω_{ij} - соответственно весовой коэффициент; T_j - порог j -го нейрона выходного слоя.

Аналогичным образом выходное значение m -го нейрона скрытого слоя определяется, как:

$$y_i = F(S_i) \tag{9}$$

$$S_i = \sum_{k=1}^n \omega_{ki} x_k + T_i, \tag{10}$$

Алгоритм обратного распространения ошибки минимизирует среднеквадратичную ошибку нейронной сети. Для этого с целью настройки синаптических связей используется метод градиентного спуска в пространстве весовых коэффициентов и порогов нейронной сети.

Согласно методу градиентного спуска изменение весовых коэффициентов и порогов нейронной сети происходит по следующему правилу:

$$\omega_{ij}(t+1) = \omega_{ij}(t) - \alpha \frac{\partial E}{\partial \omega_{ij}(t)}, \tag{11}$$

$$T_j(t+1) = T_j(t) - \alpha \frac{\partial E}{\partial T_j(t)}, \tag{12}$$

где E - квадратичная ошибка нейронной сети для одного образа.

Она определяется, как

$$E = \frac{1}{2} \sum_{j=1}^m (y_j - b_j)^2, \tag{13}$$

где b_j - эталонное выходное значение j -го нейрона.

Применимость данной архитектуры оценивалась на распознавании цифр. На вход подавались данные, полученные в результате фильтрации по алгоритму преобразования Фурье. При этом исследовалось влияние количества элементов в скрытом слое на качество распознавания изображений. На рис. 12 приведен график зависимости процента правильно распознанных цифр от данного параметра. Из графика видно, что наилучшие результаты (около 94%) достигаются при числе нейронов в скрытом слое равном 36.

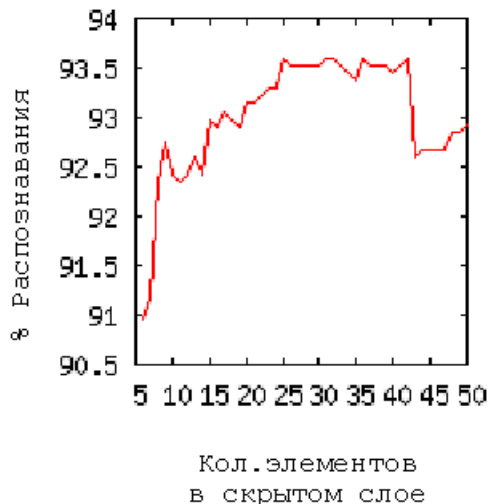


Рис. 12. График процента распознавания входной информации.

Заключение

В работе рассмотрены нейросетевые методы для автоматического распознавания номерных знаков автомобилей. При этом зашумленное изображение регистрационного номера поступает от видеокамеры на систему распознавания, которая определяет номер автомобиля. Как показали эксперименты, использование нейросетевых методов позволяет эффективно определять номер автомобиля.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Kamat V., Ganesan S. An Efficient Implementation of the Hough Transform for Detecting Vehicle License Plates Using DSP'S / Cellular Neural Network and Their Applications. Proceedings The IEEE 31st Annual 1997 International Carnahan Conference, 1997, pp 209-218.
2. Nijhuis J.A.G., Brugge M.H., Helmholt K.A. License Plate Recognition Using DTCNNs / Security Technology 1997. Proceedings 1998 Fifth IEEE International Workshop on Publish, 1998 pp 212-217
3. Rafael A., Barroso J. Number Plate Reading Using Computer Vision / Industrial Electronic 1997. ISIE'97. Proceedings of

- IEEE International Symposium on Publish. Vol. 3., 1997, pp 761-766
4. Davies E.R. Machine vision. Theory, Algorithms, Practicalities. – Academic Press, 1990
5. А.Б.Сергеенко « Цифровая обработка сигналов»; СПб.: Питер, 2003
6. Головкин В.А. Нейронные сети: обучение, организация и применение. Кн. 4: Учеб. пособие для вузов / Общая редакция А.И. Галушкина. - М.: ИПРЖР, 2001. – 256 с.

УДК 681.3

Димаков В.М.

СИСТЕМА КАРТОГРАФИИ И ПЛАНИРОВАНИЯ МАРШРУТА ДЛЯ АВТОНОМНОГО МОБИЛЬНОГО РОБОТА

1. Введение

Планирование маршрута движения для автономного мобильного робота предполагает создание искусственного интеллекта, который бы подобно человеку анализировал состояние операционной среды и принимал решение, оптимальное с точки зрения затрат для достижения цели. К сожалению, любая система планирования маршрута движения не обладает человеческими способностями в организации структуры знаний о среде, способе анализа требуемой информации и выработке необходимого решения в короткий промежуток времени. Однако существуют технологии, которые, в той или иной мере, позволяют смоделировать «осмысленные» процессы планирования маршрута движения.

В настоящее время в робототехнике четко прослеживаются два подхода в планировании маршрута движения: *тактический* и *стратегический*. Комбинированные подходы встречаются реже, что обусловлено сложностью организации модели среды, правил поведения робота и способа взаимодействия между различными компонентами навигационной системы. Мотивация здесь обусловлена в нахождении некоторого универсального метода для организации и кластеризации знаний о среде и на их основе, планирования маршрута. К сожалению, при этом прослеживается некоторая однородность решения проблемы, что, в условиях глобальной неопределенности, иногда ставит под сомнение ее эффективность. Этот вывод подтверждается достаточно простым примером о том, что интеллект высокоразвитых биологических существ не является однородной структурой, а содержит отдельные компоненты, каждый из которых обладает своими уникальными свойствами. К сожалению, научные достижения в биологии не позволяют точно сказать, каким образом осуществляется синхронизация информации между различными участками мозга высокоразвитого биологического существа. Тем не менее, сам подход в многокомпонентной организации является объектом для изучения.

Обычно задача *тактического* планировщика связана с управлением робота в особых ситуациях, таких как объезд препятствий, вычисление траектории прохода роботом двери, парковка робота, анализ критических по времени ситуаций, во избежание случайных столкновений с различными объектами. В качестве инструмента для реализации тактического планировщика могут использоваться вероятностные методы, нейронные сети, нечеткая логика. Выбор подобных средств обусловлено, прежде всего, нечеткостью данных, получаемых от сенсорных устройств, а также динамикой изменений самой среды. В случае нейронной сети, производится обучение на конкретных примерах, позволяющих ей с достаточной долей уверенности формировать правильное решение в каждой отдельной ситуации. Особенно эффективно обработка нечеткой сенсорной информации осуществляется с помощью механизма нечеткой логики, однако здесь возникает проблема выбора необходимого множества нечетких правил, покрывающих все

множество возможных решений. Отличительной особенностью тактического планировщика является отсутствие дополнительных знаний о среде, позволяющих предсказать возможные действия робота в будущем. Такое локальное поведение навигационной системы не позволяет оптимизировать маршрут движения робота для достижения цели.

В отличие от тактического планировщика, в задачу *стратегического* планировщика входит глобальное планирование всего маршрута движения. В некоторых случаях глобальное планирование осуществляется настолько подробно, что отпадает необходимость в наличии тактического планировщика. К очевидным достоинствам здесь можно отнести однозначно прогнозируемый оптимальный маршрут движения робота, однако при этом требуются значительные затраты на изменение карты среды с учетом динамичности ее характера.

Существует фактически два направления в решении задачи глобального планирования: *метрический* на базе карты размещения (occuранcy grid) и *топологический* методы. В первом случае первоначально строится вероятностная карта размещения препятствий вокруг робота на основе информации от различного рода сенсорных устройств. В результате исследования роботом всей операционной среды создается ее подробная карта, разрешение которой определяется в виде компромисса между размером всей среды и, соответственно, размером препятствий, чтобы они были достаточно точно определены на создаваемой карте. При наличии больших операционных сред поддержка карт размещения требует больших вычислительных ресурсов [1].

Среди наиболее распространенных метрических методов глобального планирования, можно выделить *вероятностное планирование* [2,3] и *подкрепляющее обучение (reinforcement learning)* [4,5,6]. В случае вероятностного планирования первоначально осуществляется построение глобальной вероятностной карты среды, которая может содержать в себе бесконечное множество измерений [2]. Построение карты осуществляется на основе локализации по методу Монте-Карло с учетом нестационарности среды во времени и размерностью $N+3$, где N – количество ориентиров среды (т.е. ее отличительных признаков). Совершенно очевидно, что с ростом ориентиров среды ее сложность обработки резко возрастает и оценивается величиной $O(N^3)$. Эта особенность построения критически ограничивает количество особенностей среды, которые могут быть запомнены. Несмотря на то, что вероятностный подход в построении карт среды достаточно хорошо изучен, вероятностное управление роботом содержит в себе множество подводных камней. Это, прежде всего, связано с большой вычислительной сложностью принятия решения в условиях глобальной неопределенности. При этом более значительные результаты были получены при замене вероятностных алгоритмов обычными, не вероятностными [7].

Другой интересный вариант решения данной задачи глобального планирования основывается на технологии под-