

А.И. САМАРИН

НИИ нейрокибернетики им. А.Б. Когана Ростовского госуниверситета,
г.Ростов-на-Дону

E-mail: samarin@krinc.ru

НЕЙРОСЕТЕВЫЕ МОДЕЛИ В ЗАДАЧАХ УПРАВЛЕНИЯ ПОВЕДЕНИЕМ РОБОТА

Аннотация

Лекция посвящена проблемам создания систем зрительного и силового восприятия роботов, обеспечивающих их целевое функционирование, а также проблемам управления поведением робота с использованием этой информации об окружении. Показывается, что в условиях замыкания обратной связи через окружение робота практически всегда отсутствует однозначная интерпретация сенсорных сигналов, без которой невозможно управление его поведением. Снятие этой неопределенности в интерпретации сенсорных сигналов может происходить только при организации активного восприятия, а, следовательно, любая даже чисто сенсорная функция обеспечивается в информационно-управляющей системе как поведенческий акт. В рамках реализации такого системного представления рассматриваются нейросетевые модели, особенности их топологии и обучения при решении задач сенсорного восприятия и целенаправленного поведения робота.

ANATOLY SAMARIN

The Department of Neural Nets and Robotics A.B. Kogan Research Institute
for Neurocybernetics, Rostov-on-Don

E-mail: samarin@krinc.ru

NEURAL NETWORK MODELS IN PROBLEMS OF ROBOT BEHAVIOR CONTROL

Abstract

The lecture is devoted to development problems for robot visual perception systems and force sensing providing a purposeful functioning. It is shown that if there is a closed feedback connection through robot's environment, practically always no simple interpretation of sensory signals exists without which robot's behavior control is impossible. The uncertainty in sensory signal inter-

pretation can be overcome only by active perception, and hence any, even purely sensory, function is considered in information control system as a behavioral action. In the frameworks of realization of such system interpretation, we consider neural network models, their architecture and learning peculiarities in solution of problems of sensory perception and purposeful robot behavior.

Введение

Представляемые материалы лекции ни в коей мере не претендуют на аналитический обзор по проблеме применения нейронных сетей в задачах робототехники. Ознакомление с нынешним состоянием проблемы в мире заинтересованные читатели могут начать с приведенного в конце лекции списка сайтов интернета. При этом хочется выразить пожелание, чтобы это ознакомление шло по возможности через призму решения проблем, затрагиваемых в лекции.

Автор излагает результаты работ, которые проводились и проводятся в НИИ нейрокибернетики Ростовского университета фактически с момента становления в нашей стране этого направления в начале шестидесятых годов. Основателем Института был один из основоположников нейрокибернетики Александр Борисович Коган. Институт нейрокибернетики открытый в 1972 году, а до этого с 1964 года проблемная лаборатория «Биофизика», был первым в мире научным учреждением, как тогда говорили, полного бионического профиля. Одно из основных направлений института формулировалось как исследование принципов и механизмов деятельности мозга с целью реализации их в технических устройствах и системах.

Фактически с самого начала эти работы носили выраженный прикладной характер, можно сослаться лишь на пример создания в начале семидесятых годов реальной, «в железе», модели сети из ста импульсных динамических нейронов для решения задач адаптивного управления движением объекта с переменной массой. К середине семидесятых годов началось бурное развитие робототехники как научного направления кибернетики. Роботу необходимы были системы сенсорного восприятия и системы адаптивного управления с использованием информации об окружении. Работая в институте бионического профиля, автор с коллегами естественно искали решение этих задач в биологически правдоподобном ключе.

Представленные в сборниках трудов конференций «Нейроинформатика» доклады по перспективным направлениям в области нейроинформатики, позволяют не останавливаться на достижениях, связанных с практическим использованием нейросетевых решений, отметив, что никакого отношения эти достижения не имеют к проблемам обработки информации и управления в робототехнике. Я имею в виду достижения с использованием ставших уже классическими парадигм многорядных нейронных сетей перцептронного типа с обучением по методу обратного распространения ошибки, сетей Хопфилда, ассоциативных структур Кохонена, сетей адаптивного резонанса Карпендера и Гроссберга и их модификаций.

Практически во всех этих моделях речь идет о задачах классификации или сводимых к ним задачах прогнозирования результата по предыдущим значениям, задачам восстановления пропущенных данных и т.д. Обучающаяся последовательность задается в виде набора примеров требуемых значений входа и выхода сети, а оценка эффективности обучения на данном этапе оценивается по степени рассогласования текущей выходной реакции нейронной сети по отношению к искомой.

Фактически это задача интерполяции функции многих переменных, когда задается совокупность точек в пространстве «вход-выход» и необходимо найти функцию, которая проходила бы через все эти точки. Оказалось, что нейросетевые алгоритмы такие задачи могут решать достаточно эффективно, то есть нейронные сети являются универсальными адаптивными интерполяторами.

При решении же задач, связанных с организацией поведения робота, при попытке формальной постановки задачи управления его поведением с использованием информации об окружении все оказывается во много раз сложнее. Насколько эти задачи сложнее, можно проиллюстрировать таким примером.

Около сорока лет назад руководители стенфордского проекта «Искусственный интеллект робота» [1] предположили, что создать компьютерную модель внешней среды и компьютерную модель робота, а затем исследовать поведение этого модельного робота в такой модельной среде очень сложно. Они решили создать реального робота с визуальными сенсорами и колесами, запустить его в реальную внешнюю среду и исследовать его так называемое сенсомоторное поведение при задаваемых целеуказаниях. При реализации этого проекта было получено много интересных решений, но главная задача – целенаправленное поведение системы во внешней среде с использованием зрения, удовлетворительно не решена

до сих пор. Хотя достижения в том, что тогда казалось очень сложным, создании виртуальных компьютерных миров и действующих в них виртуальных роботов просто впечатляющие.

В чем же проблема, на чем спотыкаются исследователи?

Прежде всего – это проблема реального времени в реальной среде. Дело здесь не только в том, что реальный мир очень сложен, а в том, что при поведении мобильного объекта в динамической среде сенсорное восприятие является результатом взаимодействия двух систем, одна из них (робот) перемещается во внешнем мире и объекты внешнего мира также могут менять свое положение. Получается как в теории относительности – только по сенсорной информации невозможно определить, что же двигалось, а если невозможно однозначно проинтерпретировать внешнюю информацию, то невозможно и организовать управление поведением, адекватное текущей ситуации. Выходом из такого положения может быть только организация активного восприятия, когда в результате достаточно длительного обучения по результатам контролируемого системой движения сенсора она может снимать эту неопределенность в интерпретации информации о внешнем мире.

Реальная среда для системы управления мобильным роботом это, как правило, произвольный мало прогнозируемый рельеф, это необходимость по внешнему виду оценивать проходимые и мало или совсем непроходимые препятствия. Здесь уместно вспомнить случай при испытаниях разработанного в ИПУ мобильного робота «Кентавр» с лазерным дальнометром. Обученный маневрированию в искусственной среде на естественном ландшафте он начинал шарахаться от высоких травинки и на поляне, по которой мог просто проехать прямо. Задача получения знаний о внешней среде, интерпретируемых в категориях поведенческих актов, совсем не тривиальна.

Для системы зрительного обеспечения поведения робота необходимо, чтобы **зрение обладало свойствами**, которые в психологии называются **объектностью и константностью восприятия**. Робот из исходного массива градуально меняющегося поля яркостей на матрице его визуального сенсора должен получить *объектное* описание сцены, а вычленение объектов из фона, когда фон может иметь такие же яркостные и текстурные характеристики, является совсем не простой задачей и практически не имеет удовлетворительных решений в сетях перцептронного типа. Во время движения робота или его визуального сенсора яркостный массив на воспринимающей поверхности сенсора меняется достаточно сложным

образом не только вследствие перемещения изображения на матрице, но и за счет изменения ракурса объектов, а также за счет изменения пространственных отношений между наблюдателем (роботом), объектом и источником освещения. Все эти причины приводят к существенному изменению яркостной картины изображения, но система информационного обеспечения в базисе зрительной модальности должна интерпретировать зрительную сцену как неизменную (*константную*). Эти свойства системы очень сложно закладывать в систему априорно не только потому, что это трудоемко, но и потому, что для реальных расчетов всегда не хватает данных. Если же делать модельную вычислительную систему, то в ней такие данные всегда есть, так как без этих данных невозможно создать саму виртуальную модель среды. Получается, то, что может быть реализовано виртуально, часто не может быть реализовано материально.

Свойства объектности и константности восприятия должны быть присущи информационно-управляющей системе робота не только при анализе зрительной информации, но и при восприятии сигналов иной природы (слух, осязание робота). Например, для робота, снабженного датчиками развиваемых сил и моментов сил, расположенных в пальцах схвата манипулятора, в запястье и сочленениях его звеньев, показания датчиков зависят от позы манипулятора в поле сил тяготения, от места приложения и величины действующих внешних сил, от развиваемых ускорений в процессе его движения.

Показания датчиков зависят от совокупности причин, но система управления должна «точно знать», что, скажем, при переносе объекта манипулирования из точки А в точку В, хотя показания всех силовых датчиков меняются, но с объектом в схвате ничего не происходит – он не соударяется с другими объектами, не проскальзывает, не меняет свою массу и т.д. Без воспроизведения такой способности, перефразируя одно философское изречение, из-под робота, каким бы конструктивно изящным он ни был, всегда будет выглядывать бульдозер.

Может быть поэтому множество казалось бы робототехнических задач в нейросетевом базисе, связанных с распознаванием объектов внешнего мира, но решенных в статике или единственном плоском отображении, задач, решенных в классическом ключе нейрокомпьютерных технологий, когда сначала чему-то обучают на большой последовательности примеров, а потом сеть ведет себя жестко детерминированно, не находят приложения в практической робототехнике.

Теперь от реальной среды перейдем к реальному времени. Требования выполнения функций отдельных подсистем в реальном времени – это не просто желание, чтобы что-то происходило быстрее. У робота в процессе управления происходит замыкание обратной связи с использованием информации от сенсоров той или иной модальности. От временных характеристик этой обратной связи зависит устойчивость и эффективность управления. Так, например, если задержка в цепи обратной связи приводов робота Пума с шестью степенями подвижности превысит 20 мс, то наступает существенное снижение его скоростных характеристик и робот переходит в режим старт-стопного управления, образно говоря, двигается медленно с мелкой постоянной дрожью. Предельные времена задержки здесь могут быть просчитаны, исходя из динамических характеристик манипулятора.

Для канала зрительной обратной связи предельное время задержки не должно превышать 100 – 200 мс при скорости движения звеньев манипулятора или мобильного робота до 1 метра в секунду. Если время оценки зрительной ситуации будет больше, то это приведет к снижению скорости движения робота или же к переходу его в режим дискретного фрагментарного управления в цикле: посмотрел – спланировал – «закрыл глаза» – двинулся на шаг. Именно в этом режиме и двигался упомянутый ранее робот Стенфордского исследовательского института [1]. Понятно, что при таком управлении ни о каком поведении в динамически меняющейся среде не может быть и речи, так как при движении «глаза» всегда закрыты.

Конспективно остановлюсь еще на одной, пока удовлетворительно не решенной проблеме робототехники, для которой нейросетевые решения будут, скорее всего, более перспективными, так как в живых системах эта проблема решена.

Когда мы хотим, чтобы робот мог пролезть куда угодно, мог изогнуться как угодно, то его наделяют все большим количеством звеньев и сочленений для достижения этой цели. Становясь более гибким, он одновременно увеличивает свою кинематическую избыточность. При такой избыточности, как известно, прямая и обратная кинематические задачи не взаимно однозначны. При решении прямой задачи для антропоморфного манипулятора по значениям суставных углов всегда однозначно определяется положение и ориентация схвата в пространстве. Обратная же задача – по положению и ориентации схвата определить значения суставных углов, имеет множество решений.

Необходимость решения обратной задачи возникает, например, при выполнении какой-нибудь операции в режиме зрительного контроля. Зрение дает текущую и требуемую координату и ориентацию схвата в пространстве, а система управления по этой информации должна сформировать уставки на приращение суставных углов. А здесь много альтернатив и робот попадает в ситуацию буриданова осла (к какому из двух равноотстоящих стогов сена подойти, чтобы поесть).

Есть здесь и другие проблемы – организация параллельного управления звеньями манипулятора при их механической связности, организация эквивалентного управления, когда движение всех звеньев должно завершаться одновременно и так далее. Пожалуй, нейросетевые решения этих задач управления многозвенником с внешним целеуказанием наиболее полно представлены в работах Александра Алексеевича Фролова с коллегами [2]. Отмечу лишь, что попытка привести эту задачу к традиционной парадигме многоярусной сети с обучением по методу обратного распространения ошибки вынудила авторов вводить очень сложную функцию с перемножением входных переменных для отдельного нейрона, что вряд ли оправдано с позиций эквивалентности решения живому прототипу.

Заканчивая столь затянувшееся вступление, хочу подчеркнуть, что оно продиктовано не только необходимостью рассмотрения некоторых проблем нейросетевых задач в робототехнике, но и тем, чтобы показать, что современное состояние проблемы использования нейронных сетей в робототехнике напоминает ситуацию, когда что-то ищут не там где потеряли, а там где светлее.

В науке, связанной с поведением, при попытке формальной постановки существует широчайший класс задач, которые не сводятся к традиционно нейрокомпьютерным. В этих задачах критерии поведения системы задаются на одних переменных, а функции «вход – выход» ищутся на других переменных [3]. Например, элементарная для нас задача: пройти по комнате, пользуясь зрением, и не столкнуться с предметами. Здесь поведение осуществляется на переменных «от глаз к ногам», а критерий задается на тактильных переменных. Другая задача, не представляющая интеллектуальной сложности: взять красный куб. И здесь цель или критерий правильности функционирования задается на переменных цвета и формы, а реализация цели на переменных, связанных с мышечными усилиями и суставными углами руки.

Получается, что существует класс задач, решение которых в природе найдено, но с математической точки зрения они поставлены некорректно.

Здесь мы не можем использовать стандартные варианты вычисления градиентов, необходимо вводить поисковые шаги, вычислять псевдоградиенты, производить декомпозицию целей, то есть переходить к многопараметрическим задачам оптимизации с числом оптимизируемых параметров (веса связей, пороги) – десятки тысяч. Такие задачи, во-первых, практически не решаются, а, во-вторых, процедура оптимизации (обучения) в ее современной реализации выходит далеко за пределы реального времени.

Как раз задачи координированного управления поведением и активного восприятия внешней среды относятся к этому классу задач. Именно к рассмотрению таких задач мы сейчас и переходим. Многие их решения с использованием биологически правдоподобных аналогий и идей, например, концепции функциональных систем П.К. Анохина [4], приводят подчас к простым и очень изящным решениям.

Идеи и модели систем целенаправленного зрительного восприятия

Стало уже традиционным говорить, что нейροкомпьютеры – это вычислители, которые в отличие от обычных компьютеров могут решать неформально поставленные задачи, и что в отличие от обычных компьютеров способ решения этих задач не программирование, а обучение [5]. Проблемам обучения нейронных сетей посвящено столько работ, что уже кажется, если рассматривается в какой-то задаче нейронная сеть без обучения, то это уже что-то совсем не то и к нейроинформатике отношения не имеет. На самом деле это далеко не так.

Можно найти много прикладных примеров, в которых простые нейросетевые решения давали великолепный результат. Взять хотя бы нейросетевой определитель скорости перемещения робота в среде с произвольной текстурой грунта, созданный в свое время нейрокибернетиками из Вильнюсского университета [6]. Для определения пройденного пути авторы отказались от традиционного пятого пассивного колеса робота, которое вечно за что-то цепляется, и сделали систему с визуальным сенсором, смотрящим на грунт под «брюхом» робота.

Сенсор представлял собой матрицу фоторезисторов, от элементов которой шли связи на нейросетевой слой с латеральным (боковым) торможением. Известно, что сеть с латеральным торможением в зависимости от

топологических параметров (как далеко распространяется латеральное торможение) и от весов связей может использоваться как фильтр определенных пространственных частот. В описываемой работе латеральные связи организовывались в двух ортогональных направлениях и активность элементов сети суммировалась отдельно по строкам и столбцам. Столбцовый сумматор давал сигнал, пропорциональный вычленяемой пространственной частоте в одном направлении, а строковый – в перпендикулярном направлении. Авторы справедливо полагали, что в яркостной структуре изображения естественного грунта данная пространственная частота в той или иной степени всегда будет присутствовать. А, далее, когда робот движется, то с периодичностью перемещения нейросетевого фильтра-маски относительно текстуры грунта менялась и выходная активность сети. Временные интервалы между всплесками активности были пропорциональны скорости перемещения робота, давая сразу иксовую и игрековую компоненты вектора скорости. Просто и красиво.

Подобные простые нейросетевые решения будут в дальнейшем рассматриваться при анализе первичных зрительных преобразований созданного визуального препроцессора робота. Но вначале – об идеях целенаправленного восприятия.

Появление новых информационных технологий, связанных с решением задач взаимодействия автономных систем с реальным окружением (например, перспективные разработки по созданию «компьютеров реального мира» или автономных роботов, движущихся в неорганизованной среде) привело к пониманию того, что функции системы технического зрения не сводятся лишь к решению задач распознавания. Функционально зрительное восприятие предполагает, прежде всего, наличие гибкой и быстро перенастраиваемой системы, осуществляющей целенаправленное поведение во внешней среде [7–14].

Анализ работ в области интеллектуального восприятия сцен показывает перспективность исследования систем машинного зрения, ориентированных на активное действие. Под активным действием подразумевается не только пространственное перемещение сенсорного устройства, изменяющего ракурс, но и целенаправленная настройка структур, обрабатывающих изображения в зависимости от цели восприятия и текущей визуальной информации [7–9]. Активное зрение можно определить как избирательный сбор визуальной информации и анализ изображений, управляемый целями восприятия. Ориентация интеллектуального анализа на активное действие предполагает исследование новых методов избира-

тельной обработки изображений и моделей целенаправленного анализа сцен.

Проблема машинного зрения со свойствами активного восприятия для автоматического распознавания образов на основе управляемого анализа изображений интенсивно исследуется в США, Европе и Японии. В США исследовательские лаборатории в Рочестерском университете [7,10], Университете штата Пенсильвания [8], в университете Мэрилэнд [15], а также в Сарнофф-Центре развивают методы, направленные на решение задач трехмерного описания сцен и распознавания образов, целенаправленной сегментации и анализа изображений сцен на основе многокурсного восприятия. В Сарнофф-Центре [11,12] интенсивно исследуется метод интеллектуального восприятия на основе селективного сбора информации о сцене и анализа поступающей информации в иерархических (пирамидальных) структурах изображений с множеством уровней разрешения, добиваясь развития средств быстродействующей обработки изображений для задач управления автономными системами.

В Европе ряд лабораторий из Франции, Германии, Италии, Дании, Бельгии, Швеции и Англии, входящие в проект ESPRIT ("Vision as Process", BR 3038 / BR 7108), исследуют методы активного зрительного восприятия и целенаправленного внимания для анализа сцен и управления движением мобильных роботов.

Серия исследований по созданию «интеллектуального» визуального сенсора с фовеальным восприятием, проведенная в лабораториях: DIST (University Italy); DESV (Отдел электроники и информатики университета Валенсии, Испания); McGill University (Канада); привела к созданию в IMEC (международный центр микроэлектроники, Бельгия) новых ретиноподобных CCD камер (типа искусственная сетчатка) на базе CMOS технологии [15–21].

Под визуальным сенсором с фовеальным восприятием понимается датчик оптического потока, имеющий сегментированные поля зрения с разным разрешением: центральное поле – с высоким разрешением и периферическое – с низкой разрешающей способностью. Такая структура полей зрения характерна для организации сетчатки глаз биологических объектов. В общих чертах процесс анализа зрительной информации основан на двухэтапной схеме вычислений (возможен также параллельный вариант обработки): первоначально анализ ведется в периферическом поле зрения и далее система переводит фовеальную (центральную) часть

поля зрения в зоны периферии для детального анализа информации [25–29].

Прототип CCD-камеры (DIST) имеет следующие характеристики: рецептивные зоны CCD-матрицы располагаются на 30 концентрических окружностях, каждая из которых содержит 64 рецептивные зоны. Размер рецептивных элементов матрицы изменяется от центра к периферии так, что если центральные элементы имеют размер 30 мкм * 30 мкм, то крайние элементы на периферии имеют размер 412 мкм * 412 мкм. Общая площадь микросхемы – 11*11 мм. Частота кадровой развертки 50 кадров/с, а общий объем информации не более 2Кбайт на одном кадре. Такой CCD-чип дает хороший компромисс между разрешением и размерами зрительного поля, а также объемом передаваемых для анализа видеоданных.

Экспериментальный анализ систем машинного зрения, оснащенных такими визуальными сенсорами, выполнен в указанных лабораториях на примерах робототехнических задач, таких как автономная навигация, обнаружение и отслеживание подвижных объектов. Результаты экспериментов показали перспективность развития ретиноподобных визуальных сенсоров, обеспечивающих высокое сжатие входного потока зрительной информации и селективный анализ динамических изображений за счет кооперативного участия в анализе двух полей зрения.

В California Institute of Technology (Ch. Koch) и Rockwell International Corp. (B.Mathur) созданы нейроморфные микросхемы зрения (Neuromorphic Vision Chips) [27]. Структура обработки информации в этих микросхемах основана на принципах организации сетчатки глаз биологических систем и реализована на современной VLSI-технологии. Подобно ретине такой сенсор адаптируется к локальным изменениям яркости и временным изменениям сигналов, обнаруживает перепады яркости и детектирует движение (тактовая частота результатов обработки визуальной информации соответствует 30 Гц). Пока такие микросхемы представляют собой первые экспериментальные образцы новых «интеллектуальных» сенсоров и испытываются в задачах управления движением автономного транспортного средства на дорогах (визуальный сенсор в реальном времени обнаруживает разметку полос на дороге или края обочины и далее зрительная система организует визуальную обратную связь в контуре управления мобильным роботом).

Таким образом, современное состояние исследований в области создания развитых систем машинного зрения с искусственным интеллектом

указывает на перспективные направления работ нейрокибернетического характера, в которых интегрируются последние достижения классической области искусственного интеллекта и результаты моделирования нейронных сетей биологических систем.

Процесс анализа изображений в такой системе обеспечивается взаимодействием собственно системы зрения с подсистемой управления движениями датчика. В зависимости от цели восприятия производится настройка системы на выделение вполне определенной информации, далее каждый текущий шаг обработки изображений сопровождается целенаправленными рассматривающими движениями визуального сенсора, в результате которых осуществляется сбор новых визуальных данных для получения устойчивого решения зрительной задачи.

После обзора очень интересных работ в области создания алгоритмов и устройств активного целенаправленного зрительного восприятия приятно похвастаться, что одной из пионерских работ в этой области была наша работа [28] и что этот приоритет признал такой авторитет в нейроинформатике, как профессор Р. Хехт-Нильсон [29].

В конце семидесятых годов для модельной проверки применимости разработанных алгоритмов управления многозвенным роботом с использованием адаптивной нейросетевой структуры в НИИ нейрокибернетике был создан исследовательский стенд «глаз-рука», представленный на рис.1.

Изначально планировалось создать управляемый плоский многозвеновик с кинематической избыточностью, на конце которого располагается ориентированный маркер. Над рабочим столом с контрастными объектами должен был располагаться простенький визуальный сенсор с двумя угловыми степенями подвижности, который бы отслеживал «взглядом» целевой объект и конец манипулятора. По сигналу рассогласования координат направления взгляда нейронная сеть должна была формировать управление многозвеновиком для вывода его конца в целевую точку. Но затем получилось так, что основное внимание было уделено визуальному сенсору и проблема моделирования систем целенаправленного зрительного восприятия занимает нас уже более двадцати лет.

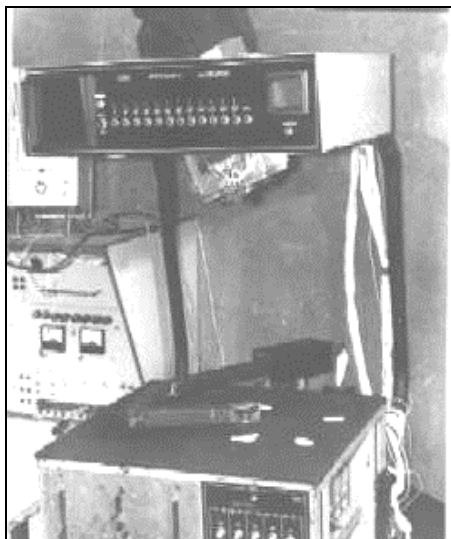


Рис. 1. Исследовательский стенд «глаз-рука» (НИИНК РГУ, 1980).

Визуальный сенсор

Основная идея сводилась к воспроизведению одного из свойств сетчатки глаза. Это свойство связано с реагированием воспринимающих элементов сетчатки только на изменение освещенности. Если изображение зафиксировать неподвижно относительно сетчатки, то через 2–3 секунды глаз перестает его воспринимать. Мы видим лишь благодаря наличию произвольных движений глаза, самое быстрое из которых – тремор, совершается с частотой 100–150 Гц и амплитудой около 20–30 угловых секунд [30]. Теперь можно представить себе такую же матрицу адаптивных воспринимающих элементов. На матрицу проецируется изображение объекта зрительной сцены, через какое-то время реакция на изображение пропадает. А теперь, если совершать случайные или циркулярные микроперемещения изображения, то будут реагировать только те элементы, на которые проецировались градуальные или контрастные перепады яркости, то есть амплитуда выходных сигналов будет пропорциональна моду-

лю градиентов яркости в локальных участках изображения. Мало того, если изображение совершает периодические циркулярные микродвижения, то в участках воспринимающей матрицы, где есть градиенты, выходные реакции элементов будут носить периодический характер. При попадании в поле зрения элементов участков изображения с разным направлением градиента (или линий контраста) эти периодические выходные сигналы будут иметь фазовые сдвиги, пропорциональные направлениям градиентов.

Схематически получающаяся картина изображена на рис.2. Аппаратная реализация такого визуального сенсора со ста параллельными каналами оптоэлектронных преобразователей представлена на рис.3.

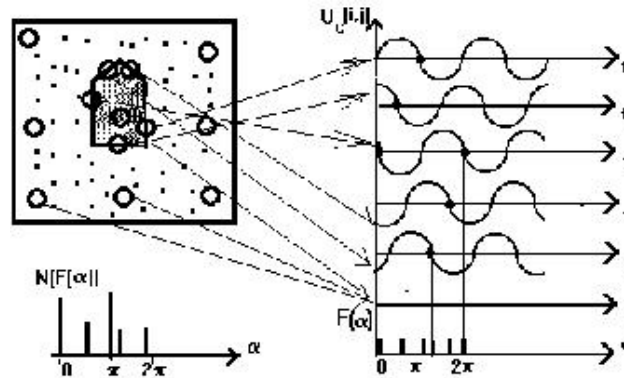


Рис.2. Схематическое представление реакций слоя первичных преобразователей (сетчатки) визуального сенсора при циркулярных микросмещениях изображения. В моменты максимальных однонаправленных приращений выходного сигнала формируются единичные импульсы, которые, суммируясь за период микросмещения, дают гистограмму распределения ориентированных перепадов $F(\alpha)$ по всему полю зрения.

Микросмещение изображения на воспринимающей матрице достигается за счет вращения эксцентрично расположенного объектива. В открытой оптической камере (рис.3) видны центральная с высоким разрешением и периферическая с низким разрешением матрицы воспринимающих элементов.

При проецировании оптического потока на такую воспринимающую матрицу распределение функции интенсивности $I(E)$ по ее элементам $E_m(i, j)$ можно записать как:

$$I(E) = \left\{ \begin{array}{l} E_m = E_m(i, j); i \in [0 \pm M]; j \in [0 \pm M]; \\ m = C(\max(i, j) \leq R); m = P(\max(i, j) > R) \end{array} \right\}; \quad (1)$$

$$\{E_C \in [0, L_C]; E_P \in [0, L_P]\};$$

где $(i=0, j=0)$ координаты центра воспринимающей матрицы; R – радиус центральной области воспринимающей матрицы, M – радиус периферической матрицы; C и P – индексы принадлежности элемента центральной или периферической матрице; L_C и L_P – количество градаций фиксируемой функции яркости в центральном и периферическом поле зрения.

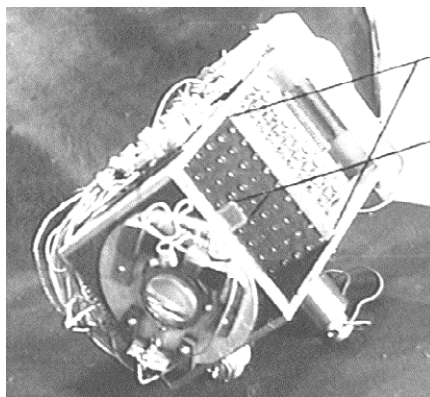


Рис.3. Визуальный сенсор с центральной и периферической воспринимающими матрицами и устройством, обеспечивающим микросмещение изображения на матрице.

В настоящей модели сенсора воспринимающие элементы центральной области матрицы реагируют на изменение интенсивности входного сигнала. В интегральном виде реакция элемента центральной матрицы $U_C(i, j)$ может быть представлена в виде:

$$E_c(t) = \int_{\Delta T} U_c(i, j) dt + (1 - \alpha) \cdot U_c(i, j); \quad (2)$$

здесь ΔT и α – параметры передаточной характеристики элемента воспринимающей матрицы.

Реакция элементов периферической матрицы описывается некоторой пороговой функцией, принимающей значение единица при превышении в зоне суммирования перепадов освещенности некоторой критической величины:

$$U_p = 1 \text{ при } \sum_s E_p(i, j) \geq \Theta; \quad (3)$$

и ноль в противном случае.

Индекс s определяет параметры зон суммирования воздействий в области периферической матрицы.

Существует механизм, обеспечивающий периодическое микросмещение изображения относительно воспринимающей матрицы. При этом каждая точка изображения движется по циркулярной орбите:

$$x = x_0 + \varepsilon \cos \omega t; \quad y = y_0 + \varepsilon \sin \omega t; \quad (4)$$

На отдельном элементе с величиной чувствительной площадки σ и центром $\mathbf{X}_0, \mathbf{Y}_0$ при угловой частоте микроперемещения ω и его амплитуде ε , не превышающей радиус чувствительной площадки каждого элемента центральной области матрицы, выходной сигнал $U_c(i, j)$ будет иметь периодическую модуляцию, если градиент освещенности $\Delta E(i, j)$ в области $\sigma(i, j)$ не равен нулю. Динамика выходного сигнала такого элемента описывается как:

$$U_c(i, j, t) = \int_0^{\infty} \frac{k}{T} e^{-\frac{\tau}{T}} \cdot E_c(i, j) [x(t - \tau), y(t - \tau)] d\tau \quad (5)$$

С учетом уравнений микроперемещения (4) для выходной функции чувствительной площадки воспринимающего элемента с координатами (i, j) выражение имеет вид:

$$U_c(i, j, t) = kE_c(i, j) - kT\varepsilon\omega \left(\frac{\partial E}{\partial x} \cdot \cos \omega t - \frac{\partial E}{\partial y} \cdot \sin \omega t \right) \quad (6)$$

Учитывая передаточную характеристику, получаемую из (2), а также то

обстоятельство, что $\sqrt{\left(\frac{\partial E}{\partial x}\right)^2 + \left(\frac{\partial E}{\partial y}\right)^2}$ – есть модуль градиента освещенности, а вектор $(\cos \alpha, \sin \alpha)$ – направляющий вектор градиента, в стационарном режиме выражение (6) принимает вид:

$$U_c(i, j, t) = K \cdot \text{mod}(\nabla E_c(i, j)) \cdot \cos(\omega t + \alpha) \quad (7)$$

Это означает, что в установившемся режиме выходные переменные элементов центральной матрицы имеют характер гармонических сигналов с периодом, равным периоду микроперемещений изображения, амплитудой, пропорциональной модулю градиентов яркости микроучастков изображения, попадающих на чувствительные площадки элементов центральной воспринимающей матрицы и фазовым сдвигом, пропорциональным направлению градиента α каждого микроучастка изображения или же направлению перепада яркости $(\alpha + \pi/2)$ этого участка.

Изменения выходного сигнала максимальны при совпадении направления микросмещения с направлением градиента перепада. И, если далее в тракте преобразования сигнала существует механизм, подобный синхронному детектированию, т.е. выделяющий максимальные перепады в различные фазы микросмещений, то фактически в такой системе на уровне первичного восприятия происходит переход в градиентное пространство признаков.

Здесь элементарным признаком выступает сверхпороговый локальный перепад яркости на изображении, ориентация которого кодирована фазой микросмещения.

Выходные переменные центральной воспринимающей области визуального сенсора представляются в виде импульсных величин единичной амплитуды в моменты максимального приращения сигнала $U_c(i, j, t)$, то есть:

$$F_{i,j}(\alpha) = \delta((t - t_0)/2\pi); \quad t_0 = t(\dot{U}_c(ij) = \dot{U}_{\max}); \quad \alpha = 2\pi \cdot \frac{t_0}{T}; \quad (8)$$

или же сигналы $U_C(i, j, t)$ суммируются по некоторой локальной области и сигнал с выхода такого фильтра векторного суммирования (9) подвергается дельта-преобразованию (10):

$$U_C(i \pm r, j \pm r, t) = K \cdot \text{mod} \left(\sum_{i \pm r} \sum_{j \pm r} \nabla E_C(i, j) \right) \cdot \text{Cos} \left(\alpha + \sum_{ij} \alpha_{ij} \right) \quad (9)$$

$$F_{i,j} \left(\sum_{ij} \vec{\alpha}_{ij} \right) = \delta((t - t_0)/2\pi) [t_0 = t(\max U_C(i \pm r, j \pm r))] \quad (10)$$

В итоге в центральной зоне модели визуального сенсора происходит преобразование проецируемого яркостного изображения в массив градиентных признаков, представленный в виде периодически повторяющегося пространственно-временного паттерна единичных импульсов. Наличие импульса на выходе элемента (i, j) свидетельствует о присутствии в его чувствительной зоне сверхпорогового градиента яркости, а его фазовый или временной сдвиг относительно нулевой фазы паттерна кодирует значение направления этого градиента.

Преобразование функции яркости в поле направлений градиентов яркости может быть выполнено также путем цифровой обработки дискретных отсчетов растрового изображения с помощью одного из локальных дифференциальных операторов [31]. Для этой цели исходное изображение обрабатывается дифференциальным оператором, вид которого определяется, как правило, быстротой вычислений и простотой технической реализации, например:

$$\frac{\partial E(i, j)}{\partial x} = \sum_{k=-1}^1 E(i+1, j+k) - \sum_{k=-1}^1 E(i-1, j+k)$$

$$\frac{\partial E(i, j)}{\partial x} = \sum_{k=-1}^1 E(i+1, j+k) - \sum_{k=-1}^1 E(i-1, j+k) \quad (11)$$

Визуальный сенсор, реализующий такие цифровые алгоритмы перехода от яркостного изображения к полю ориентаций сверхпороговых гра-

диентов, был создан на базе твердотельной телекамеры, закрепленной на антропоморфном манипуляторе (рис.4), и специализированного визуального препроцессора.

Последующая обработка таких кодированных по значениям модуля и ориентации перепадов освещенности заключается в пороговой фильтрации модулей градиентов и пространственном суммировании значений одноименных ориентаций, т.е. это фактически функции отдельных нейронов.

Гистограммы ориентированных перепадов и иерархическая организация рецептивных полей

Описание интегральных характеристик зрительного образа в виде гистограммы распределения ориентаций локальных сверхпороговых перепадов является очень удобной характеристикой. Самым полезным ее свойством является инвариантность фазовых отношений мод гистограммы к положению объекта в поле зрения, его ориентации и масштабному преобразованию.



Рис.4. Общий вид исследовательского стенда с растровым визуальным сенсором, закрепленным на конце антропоморфного манипулятора.

На рис.5 представлена иллюстрация инвариантности фазовых отношений для контрастного пятиугольника. Соответствующая гистограмма распределения ориентаций имеет пять мод, амплитуды которых пропорциональны длинам сторон прямоугольника. По такой сжатой интегральной характеристике можно уже оценить класс объекта в поле зрения, его ориентацию (по общему фазовому сдвигу относительно эталонной ориентации) и, если гистограмму не нормировать, то и величину объекта.

Нужно отметить, что гистограмма будет неизменной и при нелинейных трансформациях объектов, для которых выполняется условие сохранения в образе того же самого количества таких же ориентированных локальных перепадов яркости. Объект по форме получается совсем другой, а интегральная гистограмма старая. Для того, чтобы в условиях этой сверх инвариантности все же отличать один объект от другого необходимо производить суммирование одинаково ориентированные перепадов сначала на небольших участках – рецептивных полях, а потом полученные локальные гистограммы, привязанные к координатам рецептивных полей, суммировать в рамках иерархически организованной структуры полей.

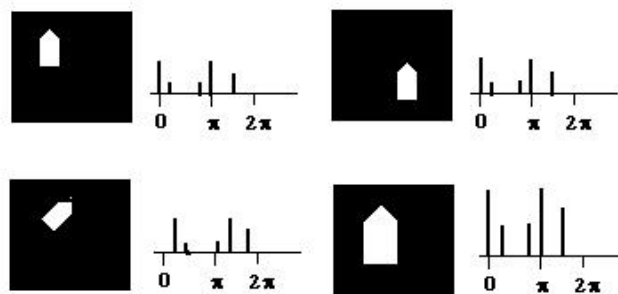


Рис. 5. Инвариантность фазовых отношений мод гистограммы при линейных трансформациях объекта в поле зрения.

Иллюстрация такой иерархической организации рецептивных полей визуального сенсора приведена на рис.6А. На рис.6Б приведен пример, показывающий, что нелинейная трансформация квадрата в крест или ступеньку при неизменности удельного веса каждой из ориентаций линий

контраста дает одну и ту же интегральную гистограмму. Но если спуститься в иерархии полей на уровень ниже (все поле зрения разбивается на 4×4 локальных рецептивных полей), то уже по локальным гистограммам объекты легко отличаются друг от друга.

Иерархическое построение рецептивных полей с разбиением дерева 4×4 (может быть и квадродерево) и представление в вершинах разбиения соответствующих гистограмм ориентированных перепадов яркости позволяет широко использовать при распознавании воспринимаемых объектов различные варианты рекурсивных алгоритмов [32].

В этом случае определяется мера близости нормированной гистограммы ориентированных перепадов яркости рассматриваемого объекта $F(\alpha) = \{\alpha_1, \alpha_2, \dots, \alpha_{16}\}$; и гистограммы эталонного объекта $F^*(\beta) = \{\beta^*_1, \beta^*_2, \dots, \beta^*_{16}\}$ при циклических сдвигах эталона на шаг $-j$. $\sum \alpha_i = 1$; α_i – нормированные значения мод гистограммы ориентаций перепадов яркости. Ориентации определяются с шагом 22.5° , весь диапазон значений ориентации перепадов яркости от 0 до 2π разбивается на 16 значений. Образ эталона виртуально поворачивается на угол j .

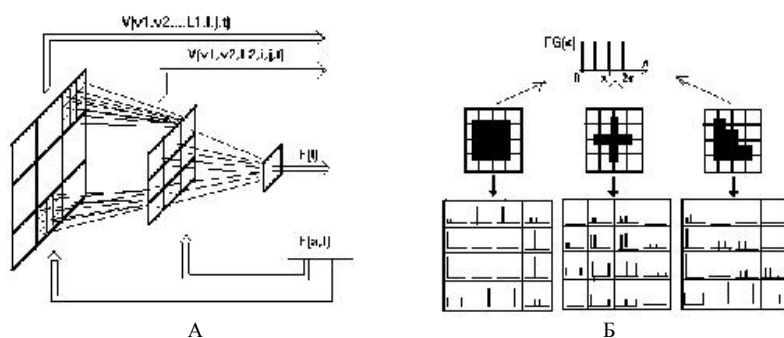


Рис.6. Иерархическое разбиение поля зрения визуального сенсора на рецептивные поля – (А), и пример уточнения описаний разных объектов с одинаковой интегральной гистограммой $F(\alpha)$ – (Б).

Определяется значение j , при котором мера близости Q максимальна:

$$Q = \max_j \left(\sum_{i=1}^{16} \min(\alpha_i, \beta_k^*) \right)_k;$$

$$k = i+j \text{ при } i+j < 16; \text{ и } k = i+j-16 \text{ при } i+j \geq 16 \\ j \in [0, \dots, 15].$$

Если величина Q превышает некоторое пороговое значение, то выдвигается гипотеза о наличии целевого объекта в поле зрения, производится разбивка поля зрения на локальные рецептивные поля с учетом найденного значения j и сопоставляются локальные гистограммы.

Пример сегментации поля зрения по присутствующим сверхпороговым градиентам и выделения целевого объекта путем оценки меры близости интегральной и локальных характеристик его описания приведен на рис.7.

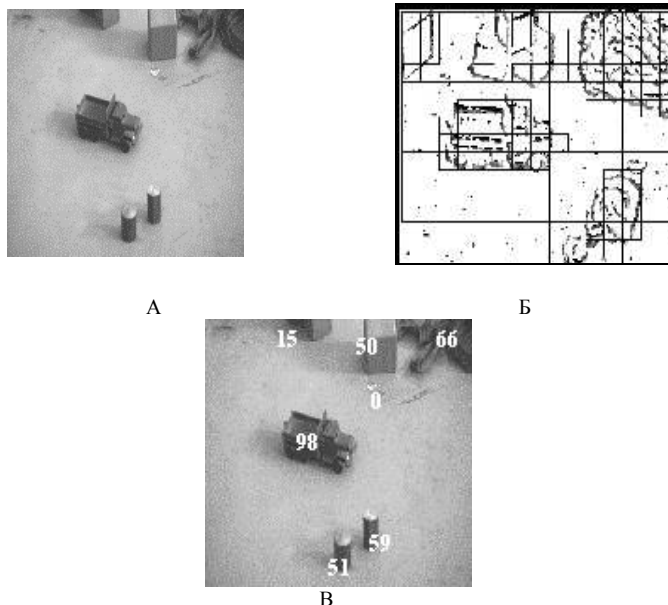


Рис.7. Пример выделения целевого объекта зрительной сцены с использованием оценки меры близости интегральной и локальных гистограмм распределения ориентированных перепадов яркости. А – исходная яркостная картина зрительной сцены, Б – сегментированные локальные поля на градиентном препарате, В – значения меры близости с целевым объектом.

Рекурсивная организация процессов анализа максимально удовлетворяет условиям задачи целенаправленного зрительного восприятия робота,

так как в зависимости от цели поведения на каждом конкретном этапе формируется конкретная цель зрительного восприятия. В такой ситуации зрительное восприятие формируется не по типу классификации образов, когда система должна ответить на вопрос «что это такое?», а по типу избирательной фильтрации, когда система отвечает на вопрос «является ли объект в поле зрения таким-то целевым объектом».

В последнем случае ответ имеет всего три значения (да, нет, не знаю). Самое главное заключается в том, что большинство отрицательных ответов может быть получено уже на основании сопоставления интегральных гистограмм. Это происходит на предельно сжатом информационном массиве и за предельно короткое время.

Обобщенная блок-схема организации процедур целенаправленного анализа в подобной системе зрительного восприятия представлена на рис.8. Хочется подчеркнуть, что система зрительного восприятия такого робота никогда не работает в режиме распознавания образов. Общепринятая схема: **сначала посмотрим, что вокруг, а потом решим, что делать**. Разрабатываемая система работает по схеме: **сначала решим, что делать, а потом определим, что для этого нужно видеть**. При такой схеме происходит очень быстрая потоковая обработка зрительной информации.

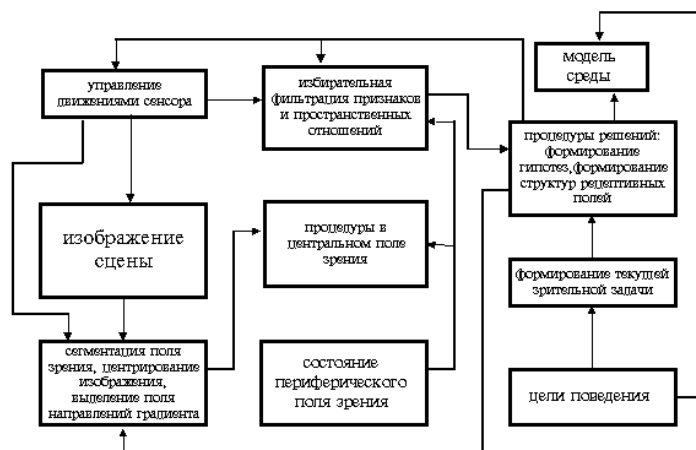


Рис.8.Схематическое представление организации процедур целенаправленного зрительного анализа.

Активное зрение и фовеальная организация рецептивных полей

При иерархической организации структур зрительного анализа разбиение на рецептивные поля никак не привязано к размеру и положению объекта в поле зрения, поэтому могут возникать случаи, когда при полном совпадении ориентационных описаний объекта на верхнем уровне, его фрагменты попадают с некоторым сдвигом в поля нижнего уровня. В этом случае локальные гистограммы рассматриваемого объекта могут достаточно сильно отличаться от таковых в эталонном описании. Поэтому приходится перемещать объект в поле зрения (смещать направление взгляда на небольшую величину) и делать сравнения с эталоном по нескольким смещенным кадрам.

Похоже, что природа выбрала несколько иной путь. Вместо иерархической организации рецептивных полей визуального сенсора на эволюционно высшей ступени развития зрительной системы появляются фовеально организованные сенсоры (фовеа – центральная часть сетчатки с высокой плотностью фоточувствительных элементов). Пространственное разрешение признаков здесь уменьшается не от уровня к уровню, а от центра сенсора к его периферии. Появляется развитая система управления движением глаза и переход на детальный анализ осуществляется не переходом на более низкий уровень интеграции рецептивных полей, а переводом взгляда (области максимального разрешения) на анализируемый фрагмент объекта. Возникает новая функция рассматривания зрительной сцены и формирование нового способа описания зрительных образов в координатах и отношениях координат взгляда в так называемых точках фиксации.

Структура рецептивных полей рассматриваемой модели фовеального сенсора представлена на рис.9. Центральное поле зрения окружено периферическим полем, которое разбито на 16 секторов. В центральном поле на выходе процессора изображений формируется интегральная гистограмма направлений градиентов яркости. В каждом из 16 секторов также формируется гистограмма направлений градиентов. В результате обработки текущего изображения фовеальный сенсор имеет на выходе 17 информационных каналов, поставляющих сжатое описание изображения.

Управляемыми величинами для такого сенсора являются координаты центра поля зрения – x_i , y_i и радиус зоны анализа (зоны внимания) – R . Активное перемещение визуального сенсора предполагает наличие двух типов движений. Первый тип движений сенсора обеспечивает направлен-

ное изменение ракурса восприятия в объемной среде, а второй тип движений управляет изменением направления точки фиксации взгляда (области максимального пространственного разрешения) в плоскости изображения.

Простейший вариант фовеальной организации, когда в поле зрения формируется зона внимания с программно изменяемыми координатами точки фиксации взгляда (центр зоны внимания) и программно изменяемым радиусом этой зоны. Вся зона разбивается на 16 радиальных секторов. В каждом из секторов производится суммирование одинаково ориентированных перепадов яркости. Все направления градиентов яркости (360°) разбиваются на 16 интервалов.

В итоге для данной точки фиксации взгляда при известном радиусе зоны внимания характеристики зрительного образа представляются в виде матрицы 16×16 значений распределения разноориентированных перепадов освещенности по секторам. Пример такого представления зрительных признаков как фрагмент процедуры рассматривания при запоминании зрительного образа приведен на рис.10.

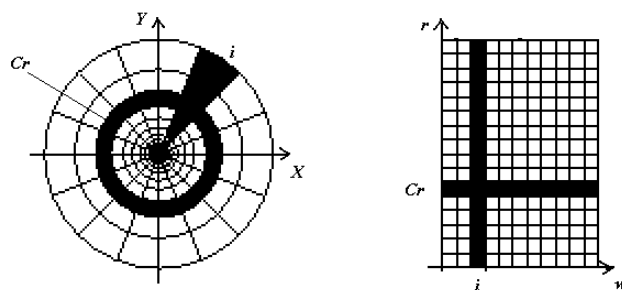


Рис. 9. Структура радиальных рецептивных полей фовеального сенсора и логарифмическое-полярное преобразование координат.

Формирование эталонной модели объекта происходит путем его осматривания по определенным алгоритмам с использованием так называемых охватывающих движений визуального сенсора. Для образа конкретного объекта в результате осмотра получают устойчивые к сдвигу, по-

вороту и масштабному преобразованию охватывающие точки фиксации взгляда.

Число этих точек в зависимости от формы объекта колеблется от 3-х до 16-ти. Запоминаются координаты этих точек, радиус зоны внимания и распределение ориентаций перепадов яркости по секторам для каждой точки. Таким образом объем памяти для запоминания эталона произвольного полутонового образа колеблется от 0.8 до 3.5 кбайт. На рис.11 представлен пример распределения точек фиксации взгляда при рассматривании объекта; белыми кружками с отметками (1),(2),(3) помечены охватывающие точки фиксации, рядом для каждой из них дано распределение ориентаций перепадов яркости по секторам – рецептивным полям, фовеального сенсора.

Таким образом, в результате процедуры «рассматривания» объекта автоматически строится его описание, используемое в дальнейшем при распознавании образов – узнавании объектов на сцене.

Процедура «узнавания» производится также с использованием рассматривающих движений сенсора и идентификацией точек фиксации взгляда по оценке меры близости распределений ориентаций для каждой из них. При узнавании используется процедура прогнозирования, когда по двум точкам фиксации взгляда прогнозируется распределение ориентаций в остальных произвольных точках фиксации по типу – «если это объект А, то посмотрев туда-то я должен увидеть то-то».

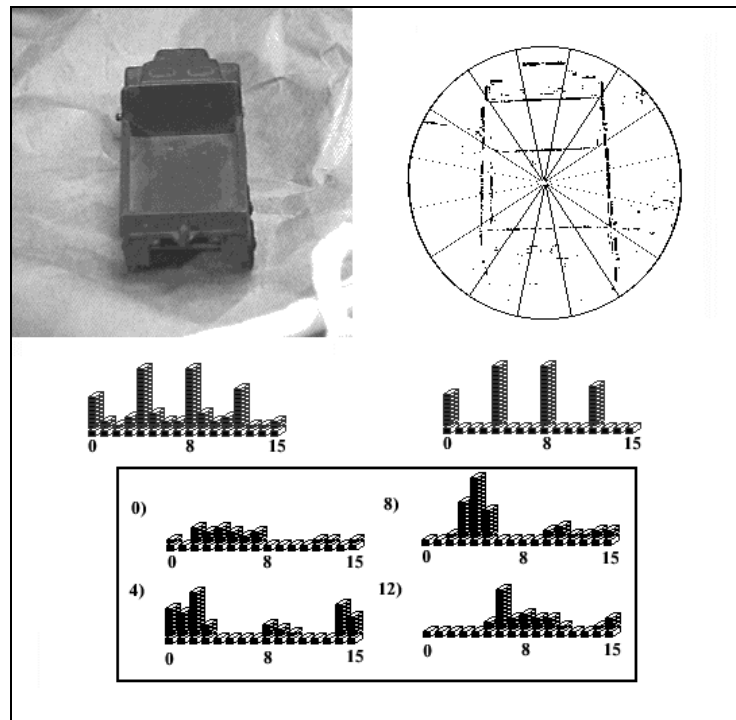


Рис. 10. Представление зрительных признаков в фовеально организованных рецептивных полях. В верхней части рисунка полутоновое изображение объекта и его сверхпороговый градиентный препарат в зоне внимания визуального сенсора; ниже – интегральная гистограмма распределения сверхпороговых ориентаций в градиентном поле изображения и гистограмма преобладающих ориентаций; в рамке – распределение преобладающих ориентаций по секторам зоны внимания.

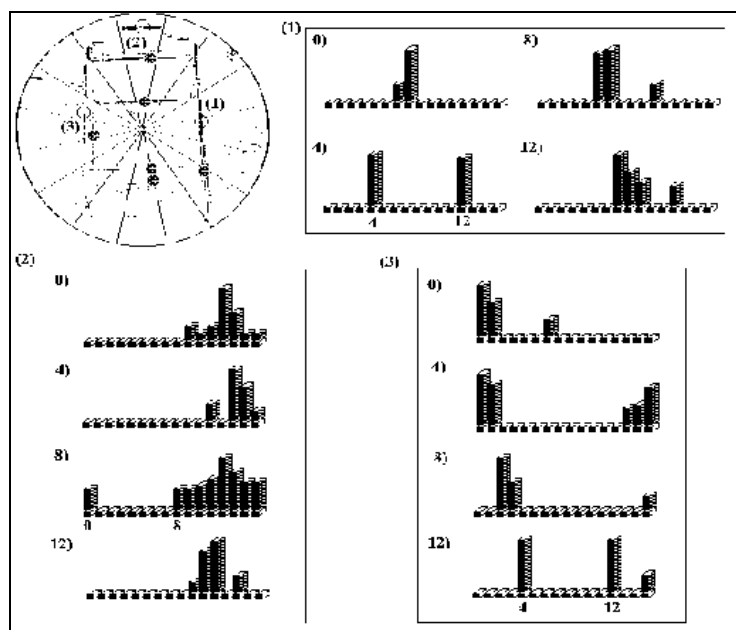


Рис.11. Распределение ориентаций сверхпороговых перепадов яркости полутонового изображения объекта по рецептивным полям сенсора для 3-х охватывающих точек фиксации взгляда.

Описанная активная система машинного зрения имеет аппаратно-программные средства поддержки и протестирована на объектах естественной полутоновой зрительной сцены.

Одно из преимуществ таких фовеальных систем заключается в том, что система координат каждый раз привязывается к характеристикам воспринимаемого объекта и результаты сопоставления интегральных и локальных характеристик образа не зависят от начального разбиения рецептивных полей. Демонстрация устойчивости этих опорных точек фиксации, относительно которых происходит привязка системы координат для последующего сопоставления локальных характеристик объектов, при трансформациях объекта в поле зрения приведена на рис.12.

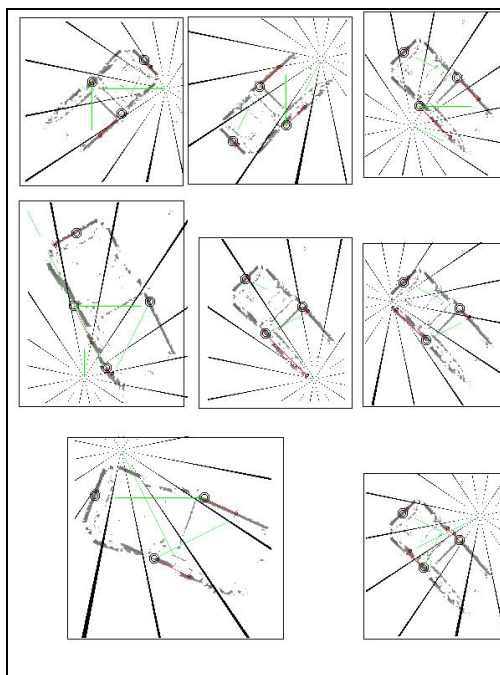


Рис.12. Пример устойчивости стабильных охватывающих точек фиксации взгляда при разных исходных позициях направления взгляда сенсора и разных трансформациях входного образа.

Теперь относительно этих охватывающих точек фиксации взгляда система зрительного восприятия может запоминать распределение признаков по радиальным полям в произвольном количестве точек. Нами показано, что по таким описаниям можно решать обратную задачу – восстановление с высокой точностью локализации пространственных признаков исходного градиентного образа. И мы можем оценивать, имеет ли система зрительного восприятия достаточное описание того или иного объекта по тому, может ли она его воспроизвести. Описанное фовеальное представление позволяет решать эту обратную задачу.

Предполагается, что фовеальную систему целенаправленного зрительного восприятия можно наиболее эффективно реализовать в нейросетевом базисе.

Нейросетевые алгоритмы распознавания зрительных образов

Большое разнообразие алгоритмов, а также широкий круг успешно решаемых задач в области распознавания образов и обработки изображений позволяет провести их некоторую классификацию и оценить с точки зрения оптимальной реализуемости в нейросетевом базисе [33].

Алгоритмы классификации или распознавания образов могут быть условно разбиты на три группы. К первой группе относятся алгоритмы, в основу которых положены идеи разделения при классификации пространства признаков с помощью секущих гиперплоскостей. Для их реализации необходима достаточно большая обучающая выборка классифицируемых образов и отсутствие единственного эталонного описания образа каждого класса. Типичным представителем этой группы алгоритмов классификации являются алгоритмы многомерного дискриминантного анализа, а классическим вариантом нейросетевой реализации являются структуры перцептронного типа [34]. Фактически функция, реализуемая отдельным нейроном, входные переменные которого определены вектором признаков распознаваемого образа, описывается уравнением плоскости в пространстве признаков, а выходная функция нейронной сети – область с разделительной гиперповерхностью, отделяющей один класс объектов от другого.

Вторую группу представляют алгоритмы распознавания, решающие правила которых построены на формировании меры близости признаковых описаний распознаваемого и эталонного образов. Естественно, если вводится мера близости или расстояние, то пространство признаков должно быть метрическим, а алгоритмы предполагают наличие эталона или объекта-представителя каждого класса образов. В какой-то степени нейросетевая реализация таких алгоритмов может представляться сетями Хопфилда [35]. При распознавании образов с помощью таких сетей предполагается, что минимумы энергетической функции сети соответствуют эталонным образам и в зависимости от расстояния между предъявляемым и каждым из эталонных образов сеть переходит в состояние к ближайшему эталону.

И, наконец, третья группа алгоритмов базируется на формировании структурно-логических описаний, чаще всего выражаемых в виде графов, вершинам которых ставятся в соответствие некоторые признаки (углы, вершины, грани), а ребра графов отображают инвариантные пространст-

венные отношения между признаками или объектами сцены [36]. Следует отметить, что для задач, связанных с анализом зрительных сцен, использование таких алгоритмов наиболее целесообразно и, по всей видимости, не имеет альтернатив в случае, когда задача строится на поиске соотношений части и целого, например, найти молоток и взять его за ручку. Но, к сожалению, нейросетевая реализация этой группы алгоритмов практически мало исследована.

Сейчас нет нейросетевых реализаций, которые позволяли бы в рамках одной и той же структуры не переобучать, а динамически поднастраивать сеть на распознавание, как в нашем примере, сначала молотка, а потом его вполне определенной части. Только удовлетворительное решение этой задачи позволит говорить об адекватных моделях механизмов зрительного восприятия живых систем. Для построения таких моделей нейронная сеть должна быть *управляемой в пространстве параметров и функционально универсальной*.

Нейросетевая реализация процесса целенаправленного анализа образов в модели активного зрения

Рассматривается модель зрительного восприятия, содержащая активный визуальный сенсор с центральным и периферическим полями зрения. Модель системы машинного зрения описанная выше и представленная в работе [37] может быть классифицирована как фовеальная система восприятия, которая выполняет пространственно-зависимую обработку изображений [38].

На первичном уровне преобразования зрительной информации осуществляется переход от яркостного к градиентному изображению. В качестве элементарных зрительных признаков система выделяет кодированные направления градиентов изображений, что позволяет формировать семантические признаки в локальных рецептивных полях в виде гистограмм распределения направлений градиентов, моды которой указывают на доминирование определенных структурных элементов возможных объектов.

Разработан метод генерации описаний образов объектов на основе формирования структурной модели в виде размеченного графа, вершины которого включают зрительные признаки различной степени сложности, а контекстные связи указывают на пространственные отношения между ними.

Описание объекта строится в результате процедуры «рассматривания» сцены визуальным сенсором с запоминанием вектора входных признаков в определенных точках фиксации взгляда. Кроме того, запоминаются пространственные отношения этих точек фиксации взгляда.

Распознавание образов производится с использованием рассматривающих движений сенсора и идентификацией точек фиксации взгляда по оценке меры близости распределений ориентаций для каждой из них. При узнавании используется процедура прогнозирования, когда по двум точкам фиксации взгляда прогнозируется распределение признаков в остальных произвольных точках фиксации с использованием условных оценок ожидания требуемых зрительных признаков в определенных рецептивных полях обрабатывающей нейронной структуры.

Нейронная сеть, реализующая эти алгоритмы, может состоять из двух модулей [39]. Один из нейросетевых модулей выполняет роль преднастраиваемого фильтра зрительных признаков, а второй модуль – сеть формирования пространственных отношений выделяемых признаков в точках фиксации взгляда.

Преднастраиваемый нейросетевой фильтр зрительных признаков имеет топологию двухслойной структуры, причем второй слой образуют *вставочные нейроны*, которые, как и в реальных нейронных структурах, осуществляют *возвратное торможение*. В этой нейронной сети связи от эфферентных нейронов на вставочные являются возбуждающими, матрица этих связей $B = (b_{ij})$, $b_{ij} > 0$; связи от вставочных нейронов к эфферентным являются тормозными, матрица связей $C = (c_{ji})$, $c_{ji} < 0$. Распределение вектора входных воздействий по эфферентным нейронам определяется матрицей $A = (a_{ki})$, $a_{ki} > 0$: i, j, k – соответственно количество эфферентных нейронов, вставочных нейронов и компонент входного вектора.

Далее, существует вторая нейронная сеть (**сеть преднастройки или сеть пространственных отношений**), которая отображает пространственные отношения выделенных фрагментов изображения в поле зрения в некоторое множество значений выходного вектора сети. Входами этой сети являются сигналы о направлении «взгляда» визуального сенсора и доминирующей ориентации перепадов яркости в центре поля зрения. Число выходов равно числу вставочных нейронов первой сети. Каждый выход второй сети заканчивается на отдельном вставочном нейроне первой сети и при активности выхода полностью тормозит связанный с ним вставочный нейрон. Это комбинаторное вытормаживающее воздействие

на вставочные нейроны может быть представлено вектор-столбцом $P = (p_1, p_2, \dots, p_j)$.

Функция нейросетевого фильтра зрительных признаков определяется как $U(t) = A * [B * C * P(t)] * F(t)$. Здесь $F(t)$ – текущий входной вектор выделяемых зрительных признаков, а $P(t)$ – текущее значение вектора пространственных отношений анализируемых фрагментов изображения. При N выходах сети преднастройки существует 2^N настраиваемых отображений $F \rightarrow U$, что уже при $N=10$ более чем достаточно для построения сколь угодно сложного описания зрительного образа. Обучаемыми в этой структуре являются веса связей, определяемые матрицами B и C , причем обучение производится отдельно для каждого значения вектора P .

Формирование выходного вектора P нейронной сети пространственных отношений осуществляется по типу обучения сети Хопфилда. Каждому новому вектору входных воздействий этой сети ставится в соответствие энергетически устойчивое состояние сети (вектор выходов P), причем хеммингово расстояние между устойчивыми значениями вектора P должно быть максимальным.

Введение условия *максимального разнообразия* выходных значений (максимальное расстояние по Хеммингу) сети пространственных отношений определяется следующими соображениями. Так как эта сеть преднастраивает в данный момент нейросетевой фильтр зрительных признаков на выделение определенного зрительного фрагмента входного образа, то желательно, чтобы обучение правильной реакции только на выделенный фрагмент для данной точки фиксации взгляда не зависело от обучения реагированию на предыдущие фрагменты в других точках фиксации взгляда. Независимость обучения может быть достигнута лишь в случае, когда для каждой точки фиксации взгляда в процедуре преднастройки участвуют (не заторможены) максимально разные вставочные нейроны. Лишь в этом случае нейросетевой фильтр зрительных признаков обучаясь адекватно реагировать на определенное распределение ориентаций зрительных признаков не разобучается правильной реакции на предыдущие входные образы.

Изложенное условие – отсутствие разобучения старого при обучении новому, которым страдают практически все алгоритмы обучения нейронных сетей с традиционной топологией, существенно сокращает время обучения нейросетевого фильтра и сводит его к коррекции возбудительных связей на активные вставочные нейроны и идущих от них тормозных связей на нейроны основного слоя.

Таким образом, сеть фильтрации обучается динамически выделять разные фрагменты изображения в зависимости от значения реакции системы на предыдущий фрагмент и текущего значения вектора направления перемещения взгляда.

Описанная модель системы целенаправленного зрительного восприятия показывает возможные пути решения выделенной во введении **проблемы объектности восприятия внешнего мира**. Другая **проблема — константность восприятия**, и ее возможные нейросетевые решения рассматривается на примере силового очувствления робота-манипулятора.

Сенсомоторное поведение робота

В качестве объекта управления рассматривается многозвенный манипулятор с кинематической избыточностью (число управляемых степеней подвижности больше числа степеней свободы). При кинематической избыточности требуемая траектория или ориентация схвата манипулятора может обеспечиваться множеством траекторий движения его звеньев или множеством его поз. Робот оснащен многокомпонентным датчиком силы в запястье и взаимодействует с объектом внешней среды. Результат силового взаимодействия робота с объектом и средой отображается в показаниях многокомпонентного силового датчика запястья.

Рассматривается задача, когда роботу дается задание переместить объект из одного положения в другое с возможным изменением его положения и ориентации в пространстве. Задаются воздействия для управления движением робота по планируемой траектории. При этом движении показания датчика запястья могут меняться очень сильно. Если с объектом в схвате ничего не происходит, то управление манипулятором не должно корректироваться по информации об изменяющихся силах. Если же объект за что-то цепляется или меняется его вес (что-то начало высыпаться) или же меняется положение какой-то его части относительно других частей, то это, естественно, влияет на показания датчика в запястье. Изменения в распределении и изменении сил при этих ситуациях могут быть значительно меньше, чем при изменении позы манипулятора в поле тяготения, но именно на эти изменения манипулятор должен реагировать. Когда информация о развиваемых силах используется при управлении положением или скоростью движения звеньев манипулятора, то такое управление осуществляется по типу искусственной податливости. Перечислены

ситуации, которые невозможно распознать с использованием дополнительных датчиков осязания – с препятствием сталкивается не манипулятор, а объект манипулирования. На выбранном примере мы сталкиваемся с классическим случаем необходимости введения свойств константности восприятия.

Чтобы выполнить это условие, робот должен перемещаться, отслеживая по динамике силового взаимодействия перемещение объекта. Так как это отслеживающее движение манипулятора связано с изменением его текущей позы или конфигурации, то получается, что при одной позе разложение вектора сил на вектор управления звеньями манипулятора будет одним, а при другой – другим. Фактически в управляющей нейронной сети нужно найти эту зависимость изменения разложения вектора сил на вектор управлений как функцию от значений суставных углов (позы манипулятора).

Найденная зависимость изменения сил при изменении позы должна отобразиться в прогнозируемых значениях вектора действующих сил силового датчика запястья. По рассогласованию текущего и прогнозируемого значений и должны формироваться корректирующие управляющие воздействия. Такой схематически описанный алгоритм достаточно полно вписывается в концепцию функциональной системы П.К. Анохина с ключевым звеном его схемы – акцептором результата действия [4].

Схематическое представление силового взаимодействия робота с внешней средой представлено на рис.13.

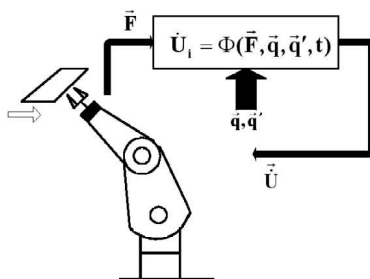


Рис.13. Манипулятор, взаимодействующий с внешним объектом, и схематическое представление адаптивного блока формирования сенсомоторной координации.

Уравнение движения робота при силовом взаимодействии может быть представлено в виде:

$$q'' = B^{-1}(q)[r(q) + \mu(t) - f(q, q') + F]$$

здесь q – вектор управляемых переменных (суставные углы); B – матрица инерции; F – вектор силового взаимодействия (сигналы с многокомпонентного датчика силы); r , f – другие факторы; $\mu(t)$ – управляющее воздействие на эффекторы ($\mu(t) = kU$).

Задача выработки сенсомоторной координации, выражается в формировании прогнозируемых значений компонент вектора измеряемых датчиком сил и в формировании корректирующих воздействий по рассогласованию реально измеряемых сил от прогнозируемых. В нейросетевой реализации ставится задача поиска функционального преобразования $U = U(F, q, q')$, при котором $Q(F) \rightarrow \min$. Здесь $U = (U_1, U_2, \dots, U_n)$ – уставки скорости изменения суставных углов q_i , n – число степеней подвижности. Нейросетевые решения этой задачи могут быть различны.

Топология нейронных сетей сенсомоторной координации

Описанная задача поиска в сети функциональных преобразований, отвечающих некоторому системному критерию, может решаться разными способами [40].

Первый, традиционный, вариант (рис.14А) заключается в построении многорядной сети с числом входов, равным числу компонент сенсорного сигнала и моторного выхода. Число слоев должно быть достаточно большим, чтобы сеть могла проинтерполировать отображаемую функциональную зависимость реальной взаимосвязи кинематики исполнительного многозвенного органа с силовыми факторами, действующими на конечное звено этого многозвенника. Такая взаимосвязь может описываться произведением тригонометрических функций, причем число компонент этого произведения не меньше числа звеньев. Отсюда – число слоев для удовлетворительного отображения искомой функциональной зависимости должно быть не меньше числа степеней подвижности манипулятора.

Наиболее важным в оценке применимости или, вернее, неприменимости традиционных вариантов обучающихся нейронных сетей является следующее соображение. В задачах, для которых традиционные нейросетевые парадигмы и алгоритмы обучения достаточно эффективны, оценка

или критерий правильного функционирования задаются как функция рас-
согласования текущего ответа сети и правильного или желаемого ответа.
Эта функция имеет унимодальный вид, обучение производится на выбор-
ке пар: входной вектор сети – выходное значение. Так как оценка унимодальна, то и процесс минимизации оценки в пространстве обучаемых па-
раметров может сводиться к использованию метода градиентного спуска с
вычислением значений частных производных оценки по подстраиваемым
параметрам сети [41].

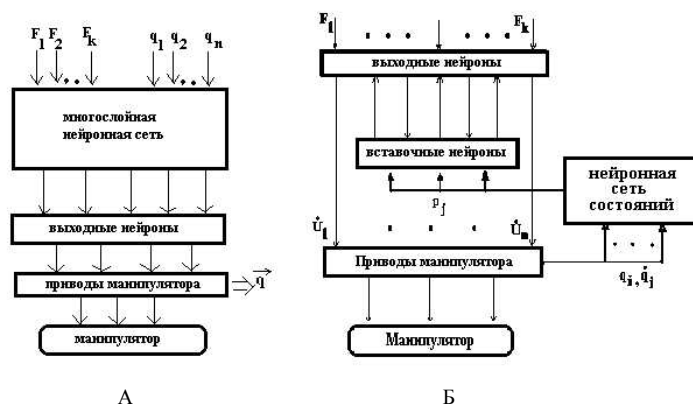


Рис.14. Два варианта топологии обучаемой нейронной сети для управления манипулятором при силовом взаимодействии с внешней средой.

А – многослойная управляющая нейронная сеть.

Б – управляющая двухслойная нейронная сеть, преднастраиваемая нейронной сетью состояний.

Рассматриваемая задача нейросетевой реализации механизмов сенсомоторной координации, к сожалению, не сводится к традиционной. Во-первых, отсутствует обучающая выборка, во-вторых, оценка задается не на расхождении текущего выхода сети от желаемого, а на входных – сенсорных, переменных, которые зависят от выхода опосредованно через исполнительные органы и внешнюю среду. А эта зависимость априорно неизвестна, для рассматриваемого примера в силу кинематической избыточности манипулятора – неоднозначна. В таком варианте функция оценки в пространстве обучаемых параметров задана в неявном виде и может иметь множество локальных минимумов.

Решение задачи предполагается следующим (рис.14Б).

Сеть сенсомоторных преобразований $F \rightarrow U$ имеет топологию двух-слойной структуры, причем второй слой образуют вставочные нейроны, которые, как и в реальных нейронных структурах, осуществляют возвратное торможение. В этой нейронной сети связи от эфферентных нейронов на вставочные являются возбуждательными, матрица этих связей $B = (b_{ij})$, $b_{ij} > 0$, связи от вставочных нейронов к эфферентным являются тормозными, матрица связей $C = (c_{ji})$, $c_{ji} < 0$. Распределение вектора входных воздействий по эфферентным нейронам определяется матрицей $A = (a_{ki})$, $a_{ki} > 0$: i, j, k – соответственно количество эфферентных нейронов, вставочных нейронов и компонент входного вектора силы.

Далее, существует вторая нейронная сеть (сеть преднастройки или сеть состояний), которая отображает в пространстве поз манипулятора некоторое множество значений выходного вектора сети. Входами этой сети являются выходы датчиков суставных углов, а число выходов равно числу вставочных нейронов первой сети (сенсомоторной сети). Каждый выход второй сети заканчивается на отдельном вставочном нейроне первой сети и при активности выхода полностью тормозит связанный с ним вставочный нейрон. Это комбинаторное вытормаживающее воздействие на вставочные нейроны может быть представлено вектор-столбцом $P = (p_1, p_2, \dots, p_j)$ с выходной функцией сети подстройки $P = P(q, t)$. Тогда функция сенсомоторной сети определяется как $U(t) = A * [B * C * P(q, t)] * F(t)$. Обучаемыми в этой структуре являются веса связей, определяемые матрицами B и C , причем обучение производится отдельно для каждого значения вектора P . Так как для такой сети частные производные критерия не вычисляются, а могут быть получены лишь их текущие оценки и существуют априорно неизвестные временные задержки между приращением аргумента и откликом критерия на это приращение, то в качестве алгоритма обучения целесообразно использовать метод многоканального синхронного детектирования со случайным опорным сигналом [41].

Сейчас мы фактически полностью повторили описание нейронной сети для реализации алгоритмов зрительного анализа при рассматривании сцен. Это не случайно. По нашему мнению, такая топология нейронных сетей максимально полно удовлетворяет условиям их применимости в задачах активного восприятия внешней среды роботами.

В целом процесс формирования сенсомоторной координации в такой сетевой структуре протекает следующим образом. Текущей позе манипулятора соответствует определенное (произвольное) состояние вектора

выходных значений нейронной сети преднастройки. Эта сеть вытормаживает часть вставочных нейронов и при взаимодействии манипулятора с внешней средой рассогласование значения вектора силы с целевым вектором запускает механизм подстройки коэффициентов сенсомоторной сети. Манипулятор при этом куда-то смещается. Если за время T правильная реакция манипулятора на сенсорное воздействие не найдена, то для нового значения позы формируется новое значение вектора выходов сети преднастройки, включаются другие вставочные нейроны и процесс повторяется. В результате такой организации процедур обучения при активном взаимодействии робота со средой одна сеть обучается адекватно целевым условиям идентифицировать состояние манипулятора, а вторая – правильно реагировать на внешние воздействия в каждом из состояний. Исследования на математической модели плоского трехзвенника показали сходимость алгоритмов обучения.

Заключение

Предложенная модель организации нейронной сети сенсомоторной координации обладает рядом положительных качеств. В частности, процедура комбинаторного вытормаживания вставочных нейронов, на которых и происходит собственно обучение, снимает проблему разобучения старому образу при обучении новому. А эта проблема характерна практически для всех обучающихся системам. В предлагаемой же модели, если старое нужно помнить, то вставочные элементы, связанные с ним, нужно вытормозить при новом обучении. Изложенные соображения о новой нейросетевой парадигме формирования сенсомоторной координации могут быть достаточно продуктивными при синтезе других систем с активным поведением.

И, напоследок, хочется еще раз подчеркнуть, что реализация свойств объектности и константности восприятия робота возможна только на нейросетевой базе, и что без воплощения этих свойств робот будет продолжать оставаться игрушкой, а не полноценным субъектом взаимодействия с внешней средой.

Литература

1. *Нильсон Н.* Мобильный автомат, построенный с использованием принципов искусственного интеллекта // В кн.: Интегральные роботы. — М.: Мир. 1973. — с.21—40.
2. *Frolov A.A., Roschin V.Y., Birjukova E.V.* Adaptive neural network model of multijoint movement by working point velocity // *Neural Network World*, **4** (4), 1994, pp.493—508.
3. *Самарин А.И.* Перспективы нейрокомпьютерных решений различных классов задач. // В кн.: Научная сессия МИФИ-99. Всероссийская научно-техническая конференция Нейроинформатика-99. Дискуссия о нейрокомпьютерах. —М.: МИФИ, 2000. — с.72—75.
4. *Анохин П.К.* Принципиальные вопросы общей теории функциональных систем // В кн.: Принципы системной организации функций. — М.: Наука. 1973. — с.5—61.
5. *Горбань А.Н.* Обучение нейронных сетей. — М.: СП ПараГраф, 1990. — 160 с.
6. *Дамиенайтис А.К.* Пространственно-временные преобразователи сигналов движущихся изображений для восприятия скорости движения роботов-манипуляторов // В кн.: Зрительные системы. Материалы Всесоюзного симпозиума «Зрение организмов и роботов». — Вильнюс, 1987. — с.306—317.
7. *Ballard D.H.* Animate vision // *Artificial Intelligence*. — 1991, v.48. pp.57—86.
8. *Bajcsy R.K.* Active perception // *Proc. IEEE*. — 1988, v.76, No. 8. pp.996—1005.
9. *Aloimonos Y., Weiss I., Bandyopadhyay A.* Active vision // *IJCV*. — January 1988, v.1, No.4. — pp.333—356.
10. *Rao R., Ballard D.H.* An active vision architecture based on iconic representations // *Artificial Intelligence*. 1995, v.78, No.1-2. pp.461—505.
11. *Burt P.J.* Attentional mechanisms for vision in a dynamic world // *Proc. of Intern. Conf. on Pattern Recognition*, Rome, 1988, pp.977—987.
12. *Burt P.J.* Smart sensing within a pyramid vision machine // *Proc. IEEE*. —1988, v.76. — pp.1006-1015.
13. *Kuniyoshi Y., Kita N., Rougeaux S., Shuehigo T.* Active stereo vision system with foveated wide angle lenses // *Proc. of the 2nd Asian Conf. On Computer Vision*, Singapore, 1995, 359-363.
14. *Hecht-Nielsen R., Zhou Y.T.* VARTAC: A foveal active vision ATR system // *Neural Networks*, Vol.3, No. 7/8, 1995, pp.1309—1321.
15. *Wodnicki R., Roberts G.W., Levine M.D.* A foveated image sensor in standart CMOS technology // *Proc. of the Custom Integrated Circuits Conference*, Santa Clara, May 3, 1995.
16. *Boldic M., Levine M.D.* A review of biologically-motivated space-variant data reduction models for robotic vision. Technical Report TR-CIM-95-05, Center for Intelligent Machines, McGill University, April 1995. (CVIU-Oct.1996)
17. *Boldic M., Levine M.D.* A Real-time foveated sensor with overlapping receptive fields. Technical Report TR-CIM-95-06, Center for Intelligent Machines, McGill University, March 1995. (Real-Time Imaging — May ,1996)
18. *Ferrari F., Hielsen J., Questa P., Sandini G.* Space variant imaging. // *Sensor Review*, Vol.15, issue No.2, 1995, pp.17—20.
19. *Pardo F., Boluda J.A., Perez J.J.* Design issues on CMOS space-variant image sensors // *Proc. of the SPIE Conference on Advanced Focal Plane Arrays and Electronics Cameras*, AFPAEC'96, Berlin, Oct. 1996.

20. *Tistarelli M., Sandini G.* Dynamic aspects in active vision. // CVGIP-Image Understanding, 56:108-129. Special Issue on Purposive, Qualitative, Active Vision, Y.Aloimonos (ed.), 1992, 56, pp.108–129.
21. *Tistarelli M.* Active/space variant object recognition. // Image and Vision Computing, 1995, 13(3), pp.215–226.
22. *Rao R.P.N., Zelinsky G.J., Hayhoe M.M., Ballard D.H.* Eye movements in visual cognition: A computational study / Technical Report 97.1, Department of Computer Science, University of Rochester, March 1997.
23. *Rao, R.* Top-down gaze targeting for space-variant active vision // Proc. ARPA'94, vol. II, pp.1049–1058.
24. *Sato Y., Toyoda M.* Active vision with two differentiated visual fields // Proc. ICPR'92, vol. I, pp.31–35.
25. *Gieffing G.J., Jan H., Mallot H.* Saccadic object recognition with an active vision system // Proc. ICPR'92, vol.I, pp.664–667.
26. *Shevtsova N.A., Faure A., Klepatch A.A., Podldchikova L.N., Golovan I.A., Rybak I.A.* Model of foveal visual preprocessor.// Proc. of the SPIE Symposium PHOTONICS EAST'95, Philadelphia, Pennsylvania, 1995, pp.588–597.
27. *Koch C., Mathur B.* Neuromorphic vision chips // IEEE Spectrum, 33 (5), 1996, pp.38–46.
28. *Самарин А.И.* Бионическая модель системы «глаз-рука» зрительно-двигательной координации робота // Доклады Межд. конференции по бионике, Бионика-78, М.-Л., т.1, 1978, с.147–149.
29. *Hecht-Nielsen R., Zhou Y.-T.* VARTAC: A foveal active vision ATR system // *Neural Networks*, 1995. 8, pp.1309–1321.
30. *Ярбус А.Л.* Роль движения глаз в процессе зрения. — М.: Наука, 1965. — 166 с.
31. *Абду И., Прэнтт У.* Количественный расчет детекторов контуров, основанных на подчеркивании перепадов яркости с последующим пороговым ограничением // ТИИЭР, 1979, N5, с.59–70.
32. *Александров В.В., Горский Н.Д.* Представление и обработка изображений: Рекурсивный подход. — Л.: Наука, 1985. — 192 с.
33. *Гаврилей Ю.К., Самарин А.И., Шенелев И.Е.* Преднастройка нейронных сетей в задачах зрительного анализа // В кн.: Научная сессия МИФИ-2000. 2-я Всероссийская научно-техническая конференция Нейроинформатика - 2000. Сб. научных трудов в 2-х частях. Ч1. М. МИФИ, 2000, с.161–166.
34. *Widrow B., Lehr M.A.* 30 years of adaptive neural networks: Perceptron Madaline, and backpropagation // Proceedings of the IEEE, 1990, 78(9), pp.1415–1442.
35. *Nasrabadi N.M., Li W.* Object recognition by a Hopfield neural network, SMC(21), 1991, pp.1523–1535.
36. *Бессарабов И.И., Гаврилей Ю.К., Кубатько А.Ю., Самарин А.И.* Разработка и исследование активной системы машинного зрения для управления поведением сложных объектов // В сб.: «Докл. III конф. РОАИ: новые информационные технологии», г. Нижний Новгород, 1997, с.184–188.
37. *Schwartz E.L., Greve D.N., Bonmassar G.* Space-variant active vision: Definition, overview, and examples // *Neural Networks*, vol.8, Number 7-8, 1995, pp.1297–1308.
38. *Психология машинного зрения* / Под ред. П.Уинстона. — М.: Мир. 1978. — 344с.

39. Самарин А.И. Формирование сенсомоторных отношений при активном взаимодействии автономной системы с внешней средой // В сб. научных трудов Всероссийской научно-технической конференции «Нейроинформатика – 99». Ч.2.М.: МИФИ. 1999. с.172–180.
40. Самарин А.И., Финкельштейн Б.А. Нейрокибернетический подход при синтезе «мозга» интеллектуального робота // В кн.: Самоорганизация и адаптивные информационно-управляющие системы. Изд-во Научного совета по комплексной проблеме «Кибернетика» АН СССР. — М., 1979. — с.76–88.
41. Расстригин Л.А. Системы экстремального управления. — М.: Наука, 1974. — 632с.

**Базовые Интернет-ссылки по проблеме приложения нейронных сетей
в задачах робототехники**

An Integrated Approach to Intelligent Systems (The University of Berkley):
<http://robotics.eecs.berkeley.edu/~layney/MURI/Intro/index.html>

MIT MURI Project:
<http://www.ai.mit.edu/projects/muri/index.html>

The Boston University MURI Project:
<http://cns-web.bu.edu/muri/>

Biomorphic Robotics:
<http://www.iguana-robotics.com/RobotUniverse/BiomorphicRobots.htm>

The Neurobotics Lab:
<http://neurobotics.bu.edu/robotics/index.html>

Rao R Активное зрение, динамическое восприятие (The University of Rochester, Computer science department):
<http://www.cs.rochester.edu/u/rao/>

Sandini G. The Space-variant CCD Sensor (The University of Genova, Italy):
<http://www.lira.dist.unige.it/Projects/Projects/Babybot/logpolar/ccd.html>

Wodnicki et al.'s foveated CMOS sensor:
http://www.pcmp.caltech.edu/visionchips/vision_chips/wodnicki.html

Foveated Imaging (The University of Texas):
<http://fi.cvis.psy.utexas.edu/>

The Log-Polar Transform and its Properties:
<http://www.ks.informatik.uni-kiel.de/~vok/research/lpt/lpt.html>

Vision Chips or Seeing Silicon:
http://www.eleceng.adelaide.edu.au/Groups/GAAS/Bugeye/visionchips/vision_chips/vision_chips.html

Computational Vision and Active Perception:
<http://www.bion.kth.se/ongoing/ongoing-html.html>

Active Attention Control Lab:
<http://www.etl.go.jp/etl/robotics/People/nkita/acac/extra/English/welcome.html>

KRINC/LACOS "Robot Vision II" project:
<http://www.krinc.ru/dnmmr/lacos96.html>

The Laboratory for Neural Network Modeling in Vision Research:
<http://www.voicenet.com/~rybak/nisms.html>

BMV: Behavioral model of Visual Perception and Recognition:
<http://www.voicenet.com/~rybak/vnc.html>

Анатолий Иванович САМАРИН – заведующий отделом моделирования нервных механизмов и робототехники НИИ нейрокибернетики им. А.Б.Когана Ростовского университета. Работу в области нейронных наук начал в 1965 году с изучения свойств функциональной универсальности нейронных сетей. Практически все последующие исследования в этой области носили прикладной характер: от задач адаптивного управления с помощью нейронных сетей, до задач создания систем технического зрения и адаптивного управления поведением мобильного робота. По тематике исследования опубликовано более 120 работ. Работа по созданию системы «Глаз–рука» в 1982 году отмечена золотой медалью ВДНХ. Серия работ по созданию комплексной системы идентификации личности погибших при массовых катастрофах с использованием нейросетевых технологий отмечена правительственной наградой. Главная мечта всей жизни – познать «как мозг делает мысль». Путь для воплощения этой мечты автор видит в реализации постулата: «Только создав робота, познаешь, как работает мозг».