

# ПРОБЛЕМЫ РЕАЛИЗАЦИИ ЗРИТЕЛЬНЫХ СИСТЕМ РОБОТОВ ДЛЯ НЕДЕТЕРМИНИРОВАННЫХ СРЕД

© 2010 г. А. С. Потапов, доктор техн. наук

НПК “Государственный оптический институт им. С.И. Вавилова”, Санкт-Петербург

E-mail: pas.aicv@gmail.com

Рассмотрены тенденции использования подсистем зрения в современных робототехнических системах, в частности, в бытовых роботах, функционирующих в недетерминированной среде. Указаны основные задачи обработки изображений, сопряженные с задачами навигации мобильных роботов. На примере применяемых на практике методов сопоставления изображений, полученных в замкнутом пространстве, установлены основные ограничения методов анализа изображений, используемых в мобильных роботах, и определены перспективы их дальнейшего развития.

**Ключевые слова:** зрение роботов, обработка изображений, недетерминированная среда.

Коды OCIS: 150.5758

Поступила в редакцию 17.06.2010

## Введение

Робототехника развивается достаточно давно. Устройства, которые можно условно назвать роботами, создавались тысячулетия назад. Большой всплеск интереса к автономным механизмам возник и в эпоху промышленной революции. Однако до 1930-х годов, когда в математике было formalизовано понятие алгоритма, каждая система управления каким-либо механизмом представляла собой уникальное аппаратное решение, в алгоритмическом плане зачастую весьма простое. Теория алгоритмов заложила начало такой науки об управлении, как кибернетика, а математическую суть сложных процессов управления стало возможным отдельить от их физического – биологического или технического – воплощения. Примерно с этого времени и начала зарождаться современная робототехника.

Однако потребовались еще десятилетия технического развития, чтобы роботы получили массовое применение. К первому поколению таких роботов принято относить роботов с “программным управлением” [1, с. 13], выполняющих жестко заданную последовательность операций. Естественно, программные роботы немногим превосходят роботов “докомпьютерной эпохи”.

Второе поколение – это адаптивные роботы, в которых в процессе управления используется

обратная связь от среды, наличие которой является обязательным условием гибкости поведения, его зависимости от текущей обстановки.

Термин “адаптивные роботы”, однако, недостаточно конкретен. Действительно, в теории управления часто адаптивными называются системы, хоть как-то реагирующие на информацию из внешней среды (т. е. просто очувствленные), и в то же время адаптивными могут быть названы и обучающиеся системы [2, с. 15].

Очевидно, диапазоны условий, в которых достигается адаптация тем или иным роботом, могут быть очень разными. В простейшем случае работает так называемая параметрическая адаптация, в которой на основе обратной связи от среды оптимизируется фиксированное число параметров закона управления.

К следующему поколению роботов относятся интеллектуальные роботы [3], для которых, однако, не существует четкого определения. Как правило, полагается, что в этих роботах используются различные технологии искусственного интеллекта [1, с. 15] – распознавание образов, машинное обучение, базы знаний, механизмы логического (или иного) вывода. С функциональной точки зрения от этих роботов требуется максимальная автономность в максимально недетерминированных условиях.

Очевидно, что в основе этих классификаций лежит в первую очередь степень недетермини-

рованности среды. Действительно, роботы первого поколения работают в полностью детерминированной среде. Такая среда может быть априорно смоделирована, что устраняет необходимость в получении из нее какой-либо информации в процессе функционирования робота. Если же неопределенность выражается в отсутствии информации о фиксированном числе параметров, то возникает необходимость в сенсорной подсистеме, обеспечивающей обратную связь от среды. Однако реализация такой подсистемы и методов обработки сенсорной информации может быть очень узкоспециализированной. Дальнейшее увеличение степени недетерминированности среды (при сохранении автономности робота) заставляет уже говорить об интеллектуальных роботах. Таким образом, деление роботов на поколения весьма условно, поскольку недетерминированность среды может повышаться плавно.

По-прежнему основным ограничением интеллектуальности роботов являются их возможности в обработке сенсорной (и, в первую очередь, оптической) информации, вызванные существованием “семантического разрыва” между массивами входной информации (например, яркостями пикселов на изображении) и теми терминами, в которых формулируются человеком задачи, ставящиеся перед роботами, работающими в максимально недетерминированной среде, в которой произвольно не только взаимное расположение объектов, но и сами типы объектов.

Таким образом, развитие роботов идет в направлении повышения степени недетерминированности среды, в которой осуществляется функционирование, при сохранении автономности роботов. Степень адаптивности и интеллектуальности робота напрямую зависит от эффективности подсистемы анализа поступающей информации, зачастую имеющей форму изображений. Одной из ключевых проблем, возникающих при разработке новейших роботов, оказывается проблема анализа изображений в условиях высокой априорной неопределенности.

### **Роботы, функционирующие в недетерминированной среде**

Традиционно роботы используются для выполнения работ, монотонных или связанных с опасностью для жизни людей. По сфере применения роботов можно условно разделить на следующие категории: промышленные, меди-

цинские, роботы для оборонных приложений, беспилотные летательные аппараты и роботизированные автомобили, роботы для сферы услуг, бытовые роботы.

Почти все из них, как правило, используются организациями, а не отдельными лицами. При этом либо робот функционирует в достаточно детерминированной среде (именно такие роботы обычно используются на производстве [1]), либо полностью или частично контролируется с помощью телеуправления. Последнее характерно для робототехнических систем оборонного назначения [4, 5] или роботов, используемых в медицине и космосе [6].

Иными словами, у роботов, работающих в недетерминированной среде, компенсация неопределенности осуществляется благодаря помощи человека-оператора. Часто эта помощь осуществляется косвенно, например, с помощью размещения фотограмметрических меток, уменьшающих неопределенность при навигации, однако, также существенно снижающих и адаптивность (диапазон сред) соответствующих робототехнических систем.

Кроме того, для этих целей используются специализированные или дорогостоящие системы сбора информации. Например, лазерные 3D-сканеры формируют “изображения”, непосредственно содержащие информацию о дальности, тогда как для извлечения этой информации из обычных изображений требуется преодоление существенной неопределенности. Гиперспектральные приборы позволяют получить дополнительные признаки, в пространстве которых цели разных типов обладают лучшей разделимостью, чем в признаковых пространствах, формируемых по обычным изображениям.

Конечно, использование дорогостоящих сенсоров не дает полного решения проблемы функционирования в недетерминированной среде: в конечном итоге для работы в произвольной среде все-таки необходимо повышать адаптивность и интеллектуальность роботов. Именно низкая интеллектуальность и является, по мнению многих авторов, основной причиной, по которой управляемые роботы во многих приложениях до сих пор обладают большей эффективностью [6].

В этом смысле интерес представляют бытовые роботы, которые должны быть автономными (в противном случае они заменяются обычной бытовой техникой) и в то же время не содержать чрезмерно громоздких или дорогих сенсоров (в противном случае они не будут конкурентоспо-

собны). Накладываются жесткие требования на отношение их стоимости к полезности, а также к компактности. Кроме того, условия функционирования этих роботов являются одними из наименее детерминированных. Как результат, рынок роботов этого типа стал развиваться позднее других, хотя научные исследования в этом направлении начались весьма давно. И до сих пор в этой области имеется весьма четкое разделение между бытовыми роботами, приносящими прибыли от их продаж, и роботами, служащими средством отработки новых технологий на будущее и поднятие престижа компании-разработчика. Тем не менее, уже несколько лет назад доля парка роботов вне промышленности (значительную часть которой составляют именно бытовые роботы) составляла около 50% [1, с. 15].

Можно выделить следующие классы бытовых роботов: роботы-игрушки, роботы, выполняющие работу по дому, роботы, взаимодействующие с людьми (роботы, помогающие инвалидам или пожилым людям, приглядывающие за детьми или домашними животными: роботизированные инвалидные кресла, роботы-поводыри, роботизированные люльки).

Роботы-игрушки могут иметь существенную разницу как в цене, так и в функциональных возможностях. В простейшем виде эти роботы мало чем отличаются от обычных механических игрушек, в связи с чем их стоимость может быть крайне низкой. Однако наделение такого робота какой-либо функциональностью требует от него определенной адаптивности, а значит, и обработки сенсорной информации, для чего необходимо снабжать робота "бортовым" вычислителем в дополнение к некоторым сенсорам и исполнительным механизмам. Многим известны роботы-игрушки, наделенные широкими функциональными возможностями по перемещению в пространстве и распознаванию отдельных объектов. Такие роботы появились на рубеже тысячелетий, и несмотря на достаточно высокую стоимость, которая для первой версии AIBO (*Sony*), вышедшей в 1999 г., составила 2500\$, имели относительно высокую скорость продаж (порядка 20 000 экземпляров в год).

Также в конце 1990-х годов появились в массовом производстве устройства, которые можно было бы назвать бытовыми роботами в полном смысле этого слова. Каждый из таких роботов должен был работать в своем помещении, которое априорно неизвестно, что сделало проблему их адаптивности принципиальной. При этом цена и потребительские качества были изначально

таковыми, что не могли непосредственно компенсировать стоимость разработки. Наиболее типичными автономными бытовыми роботами являются роботы-пылесосы. Первой моделью, выпущенной в 2002 г. фирмой *iRobot*, является *Roomba*. Затем появились роботы-пылесосы, выпущенные фирмами *LG*, *Samsung* и рядом других. К 2008 г. суммарные объемы продаж роботов-пылесосов в мире составили более полумиллиона экземпляров в год.

Многие ранние версии этих роботов (т. е. начала 2000-х годов) еще не были снабжены видеокамерами. Навигация осуществлялась преимущественно на основе сонаров и некоторых дополнительных вспомогательных сенсоров (например, датчиков касания). В другом относительно распространенном типе бытовых роботов – роботах-газонокосилках (например, в *Robotow*, выпущенном в 2003 г. фирмой *Friendly Robotics*) также использовались сонары. Сейчас, однако, бытовые роботы, снабженные видеокамерами и системами анализа изображений, обладают более высокой эффективностью при решении своих задач, что компенсирует их несколько более высокую стоимость и делает коммерчески более успешными. Кроме того, появляется возможность реализации функций роботов-охранников.

Роботам, помогающим людям с ограниченными возможностями, посвящено немало академических исследований, но широкого распространения они пока не получили. Видимо, это связано с высокими требованиями к надежности функционирования этих роботов и к их действительной интеллектуальности и, кроме того, с их высокой потенциальной стоимостью. К примеру, стоимость "коммуникационного" робота *Wakamaru* (*Mitsubishi*), выпущенного в 2005 г., составила 15 000\$ при отсутствии дополнительной полезной функциональности.

Таким образом, интересным примером автономных роботов, функционирующих в недетерминированной среде и существующих в количестве миллионов экземпляров, являются бытовые роботы, для которых лишь недавно стали широко применяться подсистемы технического зрения, позволяющие повысить эффективность функционирования.

## Проблемы реализации зрительных систем бытовых роботов

Для расширения целевой группы потребителей фирмы-изготовители бытовых роботов ста-

раются сделать их максимально дешевыми, что сказывается на качестве используемых сенсоров и на производительности бортовых вычислителей. В связи с этим большое внимание при разработке уделяется различного рода оптимизации: разработке дешевых, но эффективных (по отношению к решению стоящих перед роботом задач) оптических систем, высокопроизводительных алгоритмов обработки изображений и их адаптации к особенностям встраиваемых процессорных систем, для реализации которых могут использоваться цифровые сигнальные процессоры, программируемые логические интегральные схемы, а также мобильные процессоры.

В качестве оптических сенсоров в бытовых роботах чаще всего используются обычные камеры (часто с весьма низким разрешением и чувствительностью). Но нередко рассматривается возможность применения и других оптических систем. Наиболее популярными альтернативами являются так называемые “времяпролетные” [7, 8] и “панорамные” (или всенаправленные) [9, 10] камеры.

Времяпролетная камера регистрирует время прохождения света до точек сцены и формирует кадры, представляющие собой карты дальности, хотя и достаточно низкого разрешения. Информация о дальностях до точек сцены представляется весьма полезной при решении задач навигации и избегания препятствий.

Изображения, формируемые панорамными камерами, также облегчают решение ряда задач компьютерного зрения. В частности, облегчается задача построения карты помещения путем соединения изображений в связи с их большой площадью перекрывания, легче решаются задачи обзора помещений, что немаловажно для роботов-охранников. Подобные камеры могут использоваться и для построения стереоскопических панорамных систем [11]. Типичным способом реализации дешевых панорамных камер сейчас является использование зеркала с конической поверхностью (рис. 1).

Ограничность аппаратных ресурсов приводит к тому, что на бортовых вычислителях бытовых роботов оказывается возможным выполнять обработку изображений, обеспечивающую решение только основных задач, стоящих перед роботом. Базовой задачей, качество решения которой существенно улучшается благодаря использованию видеоинформации, является задача составления карты помещения с одновременной локализацией робота на этой карте (SLAM – *simultaneous localization and mapping*).

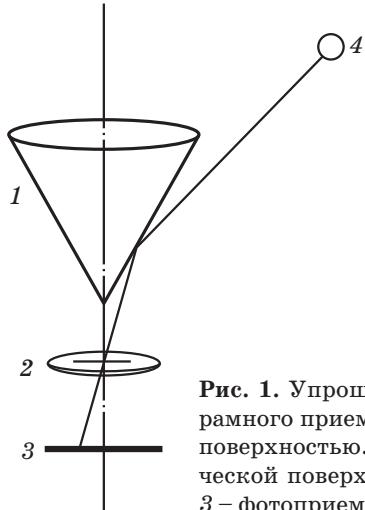


Рис. 1. Упрощенная схема панорамного приемника с конической поверхностью. 1 – зеркало с конической поверхностью, 2 – линза, 3 – фотоприемник, 4 – объект.

Действительно, большое число работ в последние годы посвящено реализации SLAM именно с помощью видеокамер [12], тогда как несколько ранее более практичным признавалось использование сонаров [13], и, как уже упоминалось, еще менее 10 лет назад большинство моделей бытовых роботов было оснащено только сонарами. Как правило, видеокамеры применяются не отдельно, а совместно с одометрами (и, возможно, другими датчиками), обеспечивающими счисление пути. При этом накапливающиеся ошибки счисления пути устраняются с помощью привязки по изображениям, а задача сопоставления изображений, полученных с разных ракурсов, упрощается за счет использования информации о приближенном положении робота, оцениваемом на основании данных одометров.

Задачи, не связанные с навигацией, но сопряженные с обработкой изображений, являются более специфическими, поскольку необходимость их решения зависит от конкретного предназначения робота. С другой стороны, нередки и задачи, которые не являются специфическими именно для бытовых роботов. К таким задачам относятся задачи обнаружения лиц, распознавания выражений лица, жестов и др., которые имеют большое значение для неробототехнических приложений (охраных систем, систем дополненной реальности и проч.), но которые характерны и для так называемых “социальных” роботов [14–16]. К примеру, способностью распознавать эмоции по выражениям лица наделен гуманоидный робот SDR-4X (*Sony*) [3, с. 303].

В отдельных случаях разработчики могут наделять свои модели роботов самыми разнообразными функциями анализа изображений, но все

же основной интерес проявляется к функциям, обеспечивающим навигацию и управление движением. Задачи навигации, для которых может использоваться обработка изображений, включают обнаружение препятствий, локализацию и картографирование [12], возврат на базу для подзарядки [17], следование за объектом [18], поиск заданного объекта [19].

В настоящее время существуют достаточно хорошо проработанные технологии решения данных задач, имеющие, однако, существенные ограничения при применении. Эти ограничения можно проследить на наиболее широко используемой технологии, которую можно было бы назвать одной из ключевых для современных систем обработки изображений в робототехнике, – технологии отождествления сопряженных точек на изображениях, полученных с разных ракурсов. Здесь чаще всего используются методы отождествления ключевых точек, или точек интереса, на основе векторов локальных признаков, вычисленных по их окрестностям. Данные методы могут использоваться не только в задачах сенсорной локализации и картографирования, но также и при распознавании объектов. К примеру, в роботе AIBO (*Sony*) использовались методы сопоставления на основе ключевых точек (разработанные фирмой *Evolution Robotics*) для распознавания предъявляемых роботу карточек (рис. 2), связываемых с разными действиями.

Для того чтобы ключевые точки могли быть попарно отождествлены, векторы признаков у сопряженных точек должны быть очень похожи (в смысле выбранной метрики, в качестве которой обычно используется евклидово расстояние в пространстве признаков). Инвариантность локальных признаков обычно достигается лишь по отношению к линейным геометрическим и яркостным преобразованиям путем аффинной нормализации окрестностей ключевых точек и нормировки самого вектора признаков. Инвариантность по отношению к несколько более сложным яркостным изменениям, вызванным изменениями условий освещения, может достигаться путем использования цветовой информации [20]. В целом, однако, несмотря на большое разнообразие методов обнаружения и описания ключевых точек, все они работают в условиях низкой изменчивости сцен и условий съемки, что заметно ограничивает возможности их применения.

Хотя используемые на практике методы анализа изображений и повышают допустимую

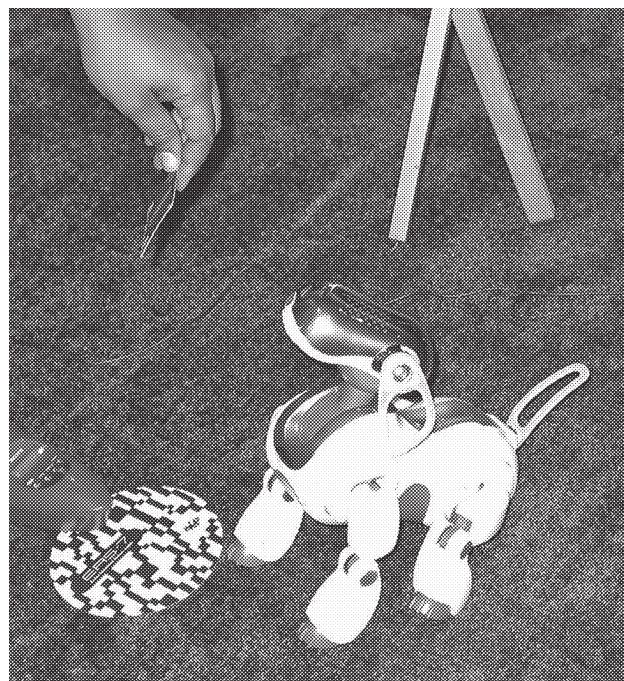


Рис. 2. Демонстрация работы системы распознавания ERSP фирмы *Evolution Robotics*, использованной в роботе AIBO (*Sony*). Фотография сделана на выставке, проходившей в 2005 году в рамках Международного симпозиума *Defense and Security*.

степень недетерминированности среды, в которой функционируют роботы, они не устраниют проблему “семантического разрыва”, препятствующую созданию интеллектуальных роботов. Сейчас расплывчатое понятие интеллектуальных роботов все чаще заменяется понятием когнитивных роботов, для которых прорабатываются более конкретные когнитивные функции, такие как внимание [21], обучение [22, 23], способы представления знаний [24] и пр., которые неразрывно связываются с механизмами зрительного восприятия. При этом все большее внимание когнитивным роботам уделяется не только в фундаментальных, но также и в прикладных исследованиях [24–27].

## Заключение

Развитие роботов следует по пути повышения степени недетерминированности среды функционирования. При этом достижение автономности функционирования роботов оказывается практически невозможным без использования подсистем зорния, включающих методы обработки изображений, работающие в условиях

высокой априорной неопределенности. Интересным примером автономных роботов являются бытовые роботы, появившиеся лишь десять лет назад, но уже сейчас выпускающиеся в общем количестве более миллиона экземпляров в год.

Наиболее проработанными технологиями, реализуемыми в масштабе реального времени на современных "бортовых" вычислителях в коммерческих мобильных роботах, являются технологии визуальной навигации и картографирования, возвращения на станцию подзарядки, избегания препятствий, распознавания небольшого числа объектов с низкой изменчивостью. При решении указанных задач еще остаются существенные проблемы, к которым можно отнести, например, сохранение работоспособности в условиях изменения освещения или при взаимных перемещениях объектов сцен. В целом, эти проблемы связаны с различными аспектами изменчивости изображений, вызванными недетерминированностью среды.

Традиционные методы анализа изображений, функционирующие в условиях низкой изменчивости сцен, не могут способствовать дальнейшему развитию автономных роботов по пути увеличения степени недетерминированности среды функционирования. Последнее требует создания интеллектуальных роботов, одними из основных компонентов которых должны быть системы понимания изображений, сопряженные с системами внимания, обучения и представления знаний.

## ЛИТЕРАТУРА

1. Юревич Е.И. Основы робототехники. 2-е изд. СПб.: БХВ-Петербург, 2007. 416 с.
2. Козлов Ю.М. Адаптация и обучение в робототехнике. М.: Наука, 1990. 248 с.
3. Bar-Cohen Y., Breazeal C. Biologically inspired intelligent robots. Washington: SPIE Press, 2003. 393 p.
4. Metcalfe J.S., Alban J., Cosenzo K.A., Johnson T., Capstick E. Field testing of tele-operation versus shared and traded control for military assets // Proc. SPIE. 2010. V. 7692. P. 769206–769206–12.
5. Cosenzo K.A., Barnes M.J. Human robot interaction research for current and future military applications: from the laboratory to the field // Proc. SPIE. 2010. V. 7692. P. 769204–769204–9.
6. Huntsberger T., Stoica A. Envisioning cognitive robots for future space exploration // Proc. SPIE. 2010. V. 7710. P. 77100D.
7. Ringbeck Th., Hagebeuker B. A 3D time of flight camera for object detection // Proc. Optical 3D

measurement techniques. Zürich. June 09–12 2007. 11 p.

8. Hsu S., Acharya S., Rafii A., New R. Performance of a Time-of-Flight Range Camera for Intelligent Vehicle Safety Applications // Advanced Microsystems for Automotive Applications. 2006. V. XVIII. P. 205–219.
9. Argyros A.A. Robot homing by exploiting panoramic vision // Autonomous Robots. 2005. № 19. P. 7–25.
10. Burbridge C., Spacek L. Omnidirectional vision simulation and robot localisation // Proc. TAROS (Towards Autonomous Robotic Systems). Guildford, UK. September 4–6 2006. P. 32–39.
11. Shih-Schön Lin. High Resolution Catadioptric Omni-Directional Stereo Sensor for Robot Vision // Proc. 2003 IEEE Intern. Conf. on Robotics & Automation. Taipei, Taiwan. September 14–19 2003. P. 1694–1699.
12. Lee Gim Hee, Marcelo H. Ang Jr. Mobile robots navigation, mapping, and localization / In Encyclopedia of artificial intelligence. N. Y.: Hershey, 2009. P. 1072–1088.
13. Leonard J.J. Directed Sonar Sensing for Mobile Robot Navigation // PhD thesis. Massachusetts Institute of Technology, Department of Engineering Science, Cambridge. 1991. 202 p.
14. Fong T., Nourbakhsh I., Dautenhahn K. A survey of socially interactive robots // Robotics and Autonomous Systems. 2003. V. 42. № 3/4. P. 143–166.
15. Breazeal C., Edsinger A., Fitzpatrick P., Scassellati B. Active vision for sociable robots // IEEE Trans. on Systems, Man, and Cybernetics. 2001. Part A. V. 31. P. 443–453.
16. Zelinsky A., Matsumoto, Heinzmann J., Newman R. Towards human friendly robots: vision-based interfaces and safe mechanisms // Lecture Notes in Control and Information Sciences. 1999. V. 250. P. 487–498.
17. Szenher M. Navigation by image-based visual homing / In Encyclopedia of artificial intelligence. N. Y.: Hershey, 2009. P. 1185–1189.
18. Comport A.I., Marchand E., Chaumette F. Robust model-based tracking for robot vision // IEEE/RSJ Intern. Conf. on Intelligent Robots and Systems, IROS'04. Sendai, Japan. September 2004. P. 692–697.
19. Tsotsos J., Shubina K. Attention and visual search: active robotic vision system that search // The 5<sup>th</sup> Intern. Conf. on Computer Vision Systems. Bielefeld, Germany. March 21–24 2007. Keynote lecture.
20. Koschan A. Improving Robot Vision By Color Information // Proc. 7<sup>th</sup> Intern. Conf. on Artificial Intelligence and Information-Control Systems of Robots. Smolenice Castle, Slovakia. September 10–14 1997. P. 247–258.

21. *Westerius C.-J.* Focus of attention and gaze control for robot vision // PhD thesis. Linkoping University, Dept. of Electrical Engineering, Sweden. 1995. 185 p.
22. *Asada M., Noda S., Tawarayamida S., Hosoda K.* Purposive behavior acquisition for a real robot by vision-based reinforcement learning // Machine Learning. 1996. № 23. P. 279–303.
23. *Martine C.M.* Genetic programming for real world robot vision // Proc. 2002 IEEE/RSJ Intern. Conf. on Intelligent Robots and Systems (IROS). Lausanne, Switzerland. September–October 2002. P. 67–72.
24. *Kira Z.* Using conceptual spaces to fuse knowledge from heterogeneous robot platforms // Proc. SPIE. 2010. V. 7710. P. 77100F.
25. *Benjamin D.P., Lyons D.* A cognitive approach to classifying perceived behaviors // Proc. SPIE. 2010. V. 7710. P. 77100H.
26. *Kelley T.D., Avery E.* A cognitive robotics system: the Symbolic and Sub-symbolic Robotic Intelligence Control System (SS-RICS) // Proc. SPIE. 2010. V. 7710. P. 77100I.
27. *Trafton G., Harrison A.* Cognitive plausible robotics // Proc. SPIE. 2010. V. 7710. P. 77100J.
-