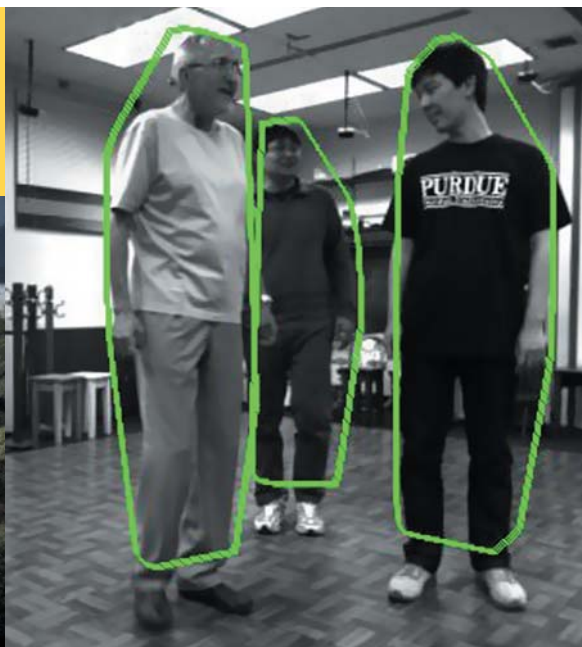




**СЕРГЕЙ КОЛЮБИН, К. Т. Н.,**

Генеральный директор  
ООО «Икстурион»,  
руководитель студенческого КБ  
по робототехнике НИУ ИТМО.



# МИР ГЛАЗАМИ РОБОТОВ

Представьте себе атлета с совершенным телом, или гонщика «Формулы 1», или могучего воина, не способных видеть. Выиграть олимпийские медали или победить врага на ощупь (голивудские блокбастеры не в счет) в этом случае им будет практически невозможно, а именно такая ситуация сложилась в современной робототехнике. Механика и электроника роботов эволюционировали настолько, что они уже превосходят человека и многие другие живые организмы и по силе, и по скорости, и по точности, и по выносливости, но пока безнадежно уступают в главном — способностях восприятия и понимания окружающей действительности.

В действительности зрение роботам нужно не только чтобы самим взаимодействовать с объектами и средой, но и чтобы быть «глазами» людей. Способность перемещаться в пространстве существенно расширяет спектр задач, в которых такие системы находят применение. Благодаря роботам инженеры могут проводить дефектоскопию многокилометровых трубопроводов или проникать в ядерные реак-

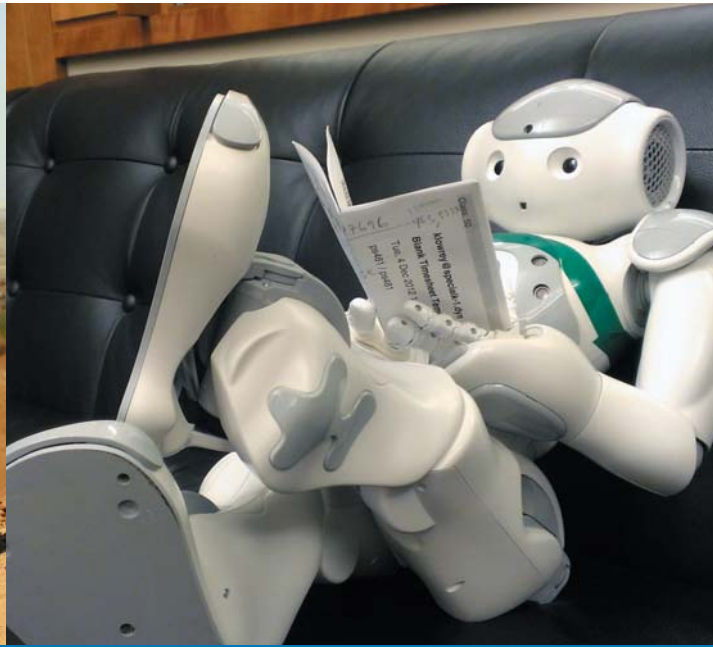
торы после аварии, как это было на «Фукусиме», а отряды полиции — разведывать ситуацию в захваченных боевиками зданиях. Роботы позволяют обследовать дно океана или запечатлеть свысока труднодоступные уголки нашей планеты, например как в уникальных роликах команды Black Sheep [1]. Марсоходы и луноходы присылают видео с далеких планет, а роботизированные хирургические комплексы позволяют погружаться в микрокосм человеческого организма. А как вам футуристические концепции квадрокоптеров для разведывания безопасного пути или предотвращения столкновения для велосипедистов и сноубордистов [2]? Конечно, от качества и надежности зрения роботов в этих задачах зависит критически много.

Как и для нас, для роботов зрение является самым богатым источником информации, но очевидно, что люди и роботы воспринимают мир по-разному. Обучить машину видеть по-человечески по сложности равносильно попытке научить человека расшифровывать в уме двоичные коды длиной в несколько километров.

По тематике машинного зрения ежегодно организуются десятки международных научных мероприятий. Только на прошедшей в ноябре прошлого года конференции по интеллектуальным роботам и системам (IEEE IROS) было представлено около 400 статей со словом «vision» в названии [3]. В них рассматриваются задачи визуальной навигации, слежения за объектами, распознавания образов и трехмерной реконструкции объектов, контроля качества в промышленности и даже визуального обучения.

Интерес к подобным задачам подогревается бизнесом. Так, в начале февраля этого года знаменитая Dyson Inc. объявила об инвестировании 5 млн фунтов стерлингов в лабораторию робототехнического зрения в Imperial College London, чтобы закончить разработку собственного интеллектуального робота-пылесоса [4]. Достаточно сказать, что на базе пары оригинальных и эффективных решений систем зрения для роботов могут выстраиваться полноценные успешные компании.

Исследовательский прогресс в этой области за последние годы



существенный, но на практике проблем по-прежнему больше, чем решений. От зрительных систем роботов требуются точность, робастность и высокая скорость работы при небольших габаритах и доступной цене. В чем же, собственно, сложность?

Сенсорика, то есть аппаратное обеспечение, развита достаточно хорошо. Отработано несколько поколений камер, включая стереокамеры, камеры со структурированным светом или ToF-камеры. Уже можно получить впечатляющие цифры мегапикселей или кадров в секунду. В подходящих условиях это гарантирует точность измерений вплоть до микронной, но алгоритмическая часть еще далека от совершенства человеческого мозга. Смещение «центра тяжести» проблемы можно проследить и по эволюции термина. Сначала использовали словосочетание «машинное зрение», потом «техническое зрение», а теперь более распространенным является понятие «компьютерное зрение».

Согласно одной из нейробиологических концепций, в зрительном центре нашего мозга выделяют области, специализирующиеся

исключительно на распознавании движений человека, животных или окружающей среды. Плюс к этому, за годы взросления мы накапливаем колоссальную базу данных зрительных образов. Соревноваться с этим пока не под силу ни одной формуле и микропроцессору. Корень проблемы в тех самых «подходящих условиях». Специалисты научились решать задачи в структурированной среде, но набегающие тени или, наоборот, яркие солнечные лучи, конвекционные потоки или перемещающиеся на фоне объекты — то, что мы в повседневной жизни даже не замечаем, — для робототехников являются серьезным вызовом. Так, успешно проходящий тесты на калифорнийских хайвеях автономный автомобиль Google может оказаться беспомощным в суровых российских реалиях отсутствующей разметки, погнутых знаков и лихачей в маршрутных такси [5]. Преодолеть это препятствие просто экстенсивным наращиванием вычислительной мощности или разрешения камер, на мой взгляд, невозможно. Нужен «умный» подход. Одним из решений может быть комплексирование информации (sensor fusion)

от системы компьютерного зрения с другими сенсорами: энкодерами, акселерометрами и гироскопами, датчиками силы, сонарами, системами глобального позиционирования. Ведь и человек опирается на разные органы чувств, а мозг принимает решения, согласовывая поступающую информацию. Мне, в частности, довелось работать в этой области в научно-исследовательском центре General Motors для проекта робота-космонавта Robonaut 2 [6], и полученные результаты внушают оптимизм. ●

#### Полезные ссылки

