

Генератор модулей распознавания для системы робототехнического зрения

А.А. Горбенко, В.Ю. Попов

*Уральский государственный университет им. А.М.Горького
gorbenko.ann@gmail.com, Vladimir.Popov@usu.ru*

Непосредственная задача систем робототехнического зрения заключается в обработке изображений, полученных с видеокамеры. При этом видеокамера может находиться непосредственно на борту робота или может быть закреплена во внешнем окружении. Такие системы позволяют распознавать объекты на изображениях, анализировать полученные данные, а затем на их основе принимать решения. Учитывая то, что, как правило, системы робототехнического зрения используются как один из ключевых элементов системы управления робототехническим комплексом (см., например, [1], [2]), спектр принимаемых ими решений, как правило, существенно шире, чем у обычных систем видеонаблюдения.

В основе любой системы робототехнического зрения лежит модуль распознавания. Он используется для выделения необходимых характеристик на изображениях. Модули распознавания могут решать задачи обработки изображений с применением различных методов и технологий. Это могут быть волновые преобразования, распознавание изображений при помощи искусственных нейронных сетей, различные алгоритмы фильтрации, сжатия и т.д. При этом обработка не обязательно состоит из какого-то одного этапа. Так, например, возможно распознавание объектов средствами только нейронной сети, заранее обученной на некоторых известных образцах. Возможно также построение распознавателя, где используется эта же нейронная сеть в сочетании с некоторым алгоритмом фильтрации. Особое значение имеют параметры нейронной сети, которые влияют на степень распознавания и количество шума, который может оставить после обработки нейронная сеть. Следует заметить, что в настоящее время существует большое количество алгоритмов фильтрации. При этом для двух различных задач может применяться одна и та же нейронная сеть, но способы фильтрации могут быть разные, или, наоборот, для различных нейронных сетей используется один алгоритм фильтрации. Исходя из этого, естественно рассматривать систему распознавания как некоторую модульную структуру, модули которой отвечают различным стадиям обработки изображения. Каждый модуль такой системы может принимать на вход некоторое изображение, в процессе работы выдавать обработанный растр, и сохранять необходимые характеристики, выявленные на этом изображении.

Создание системы распознавания для каждой конкретной элементарной задачи сводится к достаточно простой рутинной работе, требующей некоторой ручной настройки алгоритмов. Соответственно практическая реализация таких систем распознавания в большинстве случаев сводится к переписыванию одних и тех же фрагментов программного кода с небольшими модификациями. Однако полученные таким образом системы распознавания, как правило, не могут быть использованы в дальнейшем для решения каких-либо иных задач. При этом даже если мы располагаем системами распознавания для каждого из необходимых нам объектов, существенную трудность представляет механизм выбора конкретной системы распознавания в каждом конкретном случае. В то же время если попытаться усложнить исходную постановку задачи распознавания (например, вместо шара рассмотреть лицо человека), то сразу же потребуются совершенно иные технологии и иной масштаб усилий для их воплощения.

Таким образом, можно выделить три ключевые проблемы, связанные с разработкой систем распознавания: автоматизация создания простейших систем распознавания;

автоматизация механизма выбора подходящей системы распознавания; накопление и обобщение знаний для облегчения перехода от построения простейших систем распознавания к более сложным. Естественный путь для решения трех рассмотренных проблем – создание некоторой порождающей системы, способной автоматически генерировать модули распознавания по некоторому потоку визуальных данных, конструировать из этих модулей системы распознавания, а также порождать управляющие модули, обеспечивающие корректный выбор системы распознавания под конкретную задачу. При этом такая система может иметь базу знаний, хранящую информацию об уже обнаруженных и опознанных объектах.

Порождение новых модулей и систем распознавания на базе уже имеющихся представляет собой классическую задачу генетического программирования. В простейшем случае для однотипных модулей распознавания построение обобщающего модуля возможно одновременным подключением исходных модулей. Непосредственное использование одновременного подключения исходных модулей ведет к последовательному увеличению размеров порождаемых модулей, и в конечном итоге, к снижению их вычислительной эффективности. Поэтому для эффективного функционирования системы порождения необходима библиотека специализированных алгоритмов, обеспечивающих поддержание вычислительной эффективности создаваемых модулей. В частности, эта библиотека должна содержать алгоритмы минимизации; алгоритмы интеллектуального распараллеливания нейронных сетей и генетических алгоритмов; алгоритмы преобразования и встраивания методов.

Модули формирования коллективного решения представляют самостоятельную ценность для системы порождения. В частности, они могут использоваться многократно. Построение таких модулей возможно на основе генетических алгоритмов, аккумулирующих опыт ранее принятых решений. Другой эффективный метод заключается в поддержании иерархии свойств и использовании модулей, распознающих одни свойства для принятия решений по другим свойствам.

В общем случае система порождения новых модулей на базе уже имеющихся функционирует на основе аппарата регулярных выражений. При этом некоторые операции являются предопределенными. Однако вместо попытки изначально заложить некоторый максимально большой фиксированный набор удобных операций или, наоборот, ограничиться некоторым полным множеством система строится по принципу саморазвития: робототехнический комплекс на основе уже известных ему операций, применяя универсальные принципы конструирования, должен самостоятельно совершенствоваться используемый им набор операций и, в конечном итоге, логику собственного мышления.

Выполняя свою непосредственную функцию, система распознавания обрабатывает визуальный поток, извлекая из него информацию об объектах и их свойствах. Эту информацию можно использовать не только для принятия решения, но и для дополнительного обучения. В процессе обучения для каждого образа мы пытаемся сформировать набор его свойств. При этом часть из них уже может быть известна на данный момент и храниться в базе знаний. Другая часть первоначально возникает как неизвестная информация, которая может быть уточнена в последствие в результате дополнительного исследования. При этом для дополнительного исследования робототехнический комплекс не обязан ограничиваться исключительно возможностями визуальных сенсоров. Например, такая информация может быть получена в результате перемещения мобильного робота вокруг предмета. В частности, рассмотрим, например, распознавание апельсина. Когда апельсин целый, он с произвольной точки может распознаваться по свойствам «объект оранжевого цвета», «форма на растре, близкая к кругу». Если апельсин разрезать пополам и изучать его половинку, то с некоторой точки мы наблюдаем

все тот же оранжевый объект круглой формы, но в то время, когда камера перемещается вокруг половинки, на растре возникает объект оранжевого цвета в форме полукруга – это вид сбоку половинки апельсина. Далее, с противоположной исходной точки относительно апельсина мы снова будем наблюдать объект формы, близкой к кругу, но с изменившейся текстурой. Таким образом, мы получаем возможность изучить ряд новых свойств «полукруг», «апельсин в разрезе» и даже «полусфера».

Все свойства образуют иерархию. Так, например, на основе треугольника и квадрата мы можем изучить треугольную призму, как объект, который в зависимости от проекции демонстрирует на растре то свойства треугольника, то свойства квадрата. При этом в процессе обучения в общей структуре знаний появляются новые узлы, которые хранят информацию о новых свойствах или дополняют имеющиеся. Основным механизмом при создании новых свойств являются принципы обобщения и конкретизации свойств. Обобщение свойств заключается в создании свойства, родительского узла некоторого известного набора свойств. Примером может служить порождение свойств «зеленый цвет» и «круглая форма» в свойство «зеленый шар». Справедливо и обратное. Принцип конкретизации заключается в выделении новых свойств «зеленого шара»: «зеленый цвет» и «круглая форма».

Таким образом, новую ранее неизвестную информацию робототехнический комплекс преобразует в новые свойства объектов. Исследование соотношения между этими свойствами может привести к возникновению новых операций. Например, зная об отношении И, мы можем сформировать представление о новом отношении между двумя объектами на основе И, например, «между», «впереди» и т.д. В частности, в рассмотренном выше примере мы можем получить отношение «между». Если мы знаем свойства объектов «апельсин», «половинка апельсина» и «апельсин в разрезе», то отношение «между» возникает посредством движения робота вокруг половинки апельсина: свойство «половинка апельсина» проявляется «между» регистрацией свойств «апельсин» и «апельсин в разрезе».

Рассмотренный подход позволяет робототехническому комплексу порождать новые операции и отношения при помощи исследования окружения. С другой стороны, мы также можем вводить понятия операций над свойствами, используя один из универсальных механизмов порождения операций. Учитывая то, что в качестве единой платформы для модулей распознавания мы используем нейронные сети, в качестве универсального механизма порождения операций наиболее естественно использовать суперпозиции нейронных сетей, встраивая в нейронную сеть произвольной архитектуры нейронные сети, отвечающие за распознавание тех или иных свойств. С помощью нейронной сети произвольной архитектуры мы также можем вводить новые операции не на основе свойств, а на основе уже существующих операций. Если при изучении новых свойств и отношений мы используем только нейронные сети, то это позволит создать, например, новые, прежде неизвестные отношения. Но слишком большое количество таких свойств может существенно замедлить процесс обучения. База знаний в данном случае может быть переполнена различными новыми, и, возможно, похожими между собой данными. Для решения данной проблемы нам необходимо помимо использования нейронных сетей учитывать информацию, полученную во время исследования окружения. Такой подход воспрепятствует переполнению базы знаний лишней информацией и позволит ускорить обучение.

Выявляя образы, мы удаляем графический шум. В случае обучения новым явлениям или свойствам нам важна любая информация, содержащаяся на изображении. Поскольку шум – информация, не относящаяся к объектам, которые мы умеем распознавать, т.е. неизвестная для нас информация, графический шум в процессе самообучения системы играет одну из ключевых ролей. Мы должны хранить информацию о его коли-

честве и характере изменения от изображения к изображению. Анализируя данную информацию, мы можем выявить закономерности и открыть новые явления и свойства. При этом могут изменяться уже известные свойства. Собирая статистику по характеру графического шума, мы можем, также как и свойства, открывать и создавать новые явления. К ним относятся вибрации, природные явления (дождь и др.). Сбор знаний об окружающем мире заключается не только в сборе визуальной информации. В ней могут участвовать различные датчики на борту робота, способные дать дополнительную информацию о состоянии окружающей среды. По характеру работы механических частей робота совместно с визуальной информацией робототехнический комплекс также может сделать открытия новых явлений. Анализируя характер шума, система исходит из частоты и закономерности возникновения того или иного шума, а также из существенности событий, с которыми можно соотнести этот шум. Например, комнатный робот, однажды попавший во двор и оказавшийся под дождем, скорее всего отнесет дождь к несущественному белому шуму, не сумев квалифицировать его как некоторое явление. Тот же робот, всего один раз упав со стола на пол и повредив один из своих датчиков, череду "смазанных" из-за резкого перемещения кадров наверняка квалифицирует как некоторое опасное явление. Для обучения новым образам из базы знаний нам необходим дополнительный модуль. Его цель выбирать некоторые известные свойства из базы знаний и создавать новые образы, которые сочетают эти свойства. Например, наша база знаний хранит информацию о синем и красном цвете, и о двух формах – треугольник и квадрат. Данный модуль может получить, например, новый образ – синий квадрат. На основе знаний о характере изменения свойств и явлений на изображении мы можем рассмотреть задачу прогнозирования движения. В этом случае по движению некоторого образа на кадрах мы можем сделать заключение о его расположении на последующем изображении. Зная информацию о характере движения образов, мы можем выявлять новые свойства.

Наиболее интересную и важную часть системы порождения распознавателей представляет собой механизм извлечения новых знаний, не использующий ранее порожденные модули. Решение проблем подобного типа либо опирается на уже имеющуюся базу знаний, хранящую информацию об уже обнаруженных и опознанных объектах и их характеристиках, таких, как, например, цвет, размер, форма и т.д., либо осуществляется в рамках непосредственного взаимодействия робототехнического комплекса и его окружения. В последнем случае принципиальное значение имеет то, что обработка визуальной информации осуществляется не только в рамках собственно системы распознавания и генератора таких систем, для извлечения визуальных знаний задействованы все доступные робототехническому комплексу возможности.

Основные преимущества систем распознавания на основе интеллектуальных генераторов модулей распознавания и самих систем заключаются в возможности быстро адаптироваться к новым условиям функционирования и новым задачам, а также в существенном снижении затрат на создание отдельной системы при массовом использовании. Кроме того, применение таких систем дает существенный исследовательский эффект при массовом использовании и обеспечении разнообразия решаемых задач и условий работы. Исходя из этого, наибольший интерес такие системы представляют для такого бурно развивающегося направления, как сервисная робототехника (см. [3]). Во многих случаях стоимость сервисных роботов пока слишком высока, чтобы сделать экономически целесообразным их массовое использование. В последние годы ценовая характеристика сервисного робота становится одной из важнейших. Поэтому в развитых странах активно развиваются исследования, направленные на вытеснение интеллекта и соответствующей электроники с самого робота на более высокий уровень: суперкомпьютеры и даже сети суперкомпьютеров. В соответствие с этим функциониру-

вание системы распознавания на основе генератора естественно рассматривать в рамках модели «робот – управляющий компьютер – суперкомпьютер». При этом предполагается максимальное возможное снижение требований к вычислительным ресурсам роботов, подключенных к суперкомпьютеру. В рамках разрабатываемой нами системы главной единицей является генератор. Он должен иметь быстрый доступ к базе знаний, а также мощные вычислительные ресурсы, которые позволяют порождать новые модули системы распознавания; проектировать новые системы; обновлять текущую систему; по текущему изображению производить необходимые статистические расчеты для накопления знаний. В связи с этим генератор располагается на суперкомпьютере или кластере. Соответственно процесс окончательного формирования новых знаний осуществляется именно здесь. Заметим, что база знаний существенно не меняет свое содержание со временем, поэтому для предотвращения многократной пересылки этих данных и перегрузки сети, база знаний также хранится на кластере. Все модули, с помощью которых мы создаем новые варианты распознавателей также располагаются здесь. Сбор и анализ информации по текущему состоянию робота осуществляет управляющий компьютер. Он обеспечивает основные передачи данных, то есть взаимодействие управляющего компьютера с кластером или бортовым компьютером, содержит информацию о пройденных маршрутах, списке поставленных задач, статистику их выполнения и т.д. Управляющий компьютер также содержит программы для визуального представления информации. Если в системе управления используются несколько роботов, управляющий компьютер поддерживает взаимодействие между ними, передачу информации, организацию выполнения общих задач и т.д. В случае необходимости генерирования нового типа распознавателя, пакет данных посылается на кластер. Таким образом, все управление конкретным роботом сосредоточено на управляющем компьютере, кластер содержит лишь необходимую информацию, т.е. базы знаний и модули для порождения распознающих систем, и систему порождения новых знаний. Это обеспечивает надежность функционирования робота вне зависимости от наличия связи с суперкомпьютерным центром. Отметим, что в качестве управляющего компьютера может выступать не обязательно персональный компьютер, но и некоторое мобильное устройство. В этом случае возможно удаленное управление роботом или наблюдение за его состоянием. Во время удаленного сеанса мы имеем доступ к списку текущих выполняющихся задач, статистике работы различных систем робота. Периодически управляющий компьютер посылает данные на бортовой компьютер. На бортовом компьютере робота мы располагаем только два модуля. Первый модуль необходим для сбора информации от устройств. Это могут быть визуальные датчики, сенсоры и т.д. После того, как необходимый пакет данных сформирован, он передается на обработку управляющему компьютеру. Второй модуль на борту робота необходим для принятия решений в случае аварийной ситуации. Это может быть отключение компьютерной сети, разрядка аккумуляторов и т.д. Данный модуль содержит набор команд для действий в той или иной ситуации. При этом возможна реализация, в рамках которой аварийный модуль находится в архиве, а активно функционирует только анализатор, запускающий ту или иную часть аварийного модуля при необходимости.

Литература

1. P. Sala, R. Sim, A. Shokoufandeh, S. Dickinson. Landmark Selection for Vision-Based Navigation. *IEEE Trans. on Robotics*. 2006. V.22. N2. P.334-349.
2. R. Murrieta-Cid, C. Parra, M. Devy. Visual navigation in natural environments: from range and color data to a landmark-based model. *Autonomous Robots*. 2002. V.13. N2. P.143-168.
3. <http://www.robotics.org/>