

Машинное зрение: задачи и возможности С. Ю. Желтов, Ю. В. Визильтер

В последние годы цифровой анализ изображений проникает во все большее количество областей применения, таких как интеллектуальные робототехнические комплексы, системы управления движущимися аппаратами, обработка данных дистанционного зондирования, системы промышленного контроля, биомедицинские исследования, новые технологии обработки документов и пр. Термин «*машинное зрение*» (Machine vision) представляет сегодня понятие, наиболее полно охватывающее круг инженерных технологий, методов и алгоритмов, связанных с задачей интерпретации сцены наблюдения по её двумерным проекциям (изображениям), а также практическое использование результатов этой интерпретации.

Проблематика машинного зрения весьма привлекательна для современных исследователей, так как аппаратные возможности, предоставляемые в данной области последними достижениями электроники и вычислительной техники, настолько высоки, что они во многом приближаются к «техническим характеристикам» человека. Разрешение многих сенсоров для получения видеoinформации практически соответствует числу элементов сетчатки глаза человека, а возможности ЭВМ и специальных процессоров близки к характеристикам «вычислительных мощностей», используемых для обработки изображений в мозгу. Биомеханика вплотную подошла к разработке сложных механических манипуляторов, достаточных для имитации моторной деятельности человека по управлению различными техническими системами, и таким образом, на пути к осуществлению заветной мечты будущего – созданию сложных автономных робототехнических комплексов, «интеллектуальных машин» реального времени – на первый план выдвигается необходимость решения задачи разработки методов и алгоритмов «*понимания изображений*». Однако, именно это во многих случаях оказывается наиболее трудным.

Задачи машинного зрения

Удивительная сложность проблемы «понимания изображений» объясняется тем обстоятельством, что её интеллектуальная (алгоритмическая) составляющая оказалась во многом более сложной, чем традиционные задачи типа компьютерной игры в шашки или шахматы, которые долгие годы служили полем приложения методов «искусственного интеллекта». Повидимому, это связано со сложностью основного предмета, находящегося в центре внимания данной дисциплины, а именно – двумерного изображения. Скорее правилом, чем исключением, является отсутствие у информационного семантического содержания изображения какой-либо «причинной» или динамической модели формирования в том смысле, что это информационное семантическое содержание возникает не под действием каких-либо физических законов, описываемых математическими уравнениями. Информационное наполнение изображения проявляется в виде бесконечного разнообразия *яркостно-геометрических структур*, модели порождения которых могут просто отсутствовать. Особенно сложной задачей является «понимание» искусственных объектов, присутствующих в сцене наблюдения. Обнаружение и идентификация многих типов таких объектов, как здания и дороги - на аэрофотоснимках превратились в отдельные направления исследований. Проблеме выделения зданий на изображениях были посвящены несколько крупных международных конференций, поставивших лишь новые вопросы. Следует признать, что общая теория «понимания изображений» за последние 30 - 40 лет ещё не вышла из младенческого возраста. Сегодняшнее её состояние - это сочетание ряда нерешенных теоретических задач с одной стороны, и большого числа идей и подходов, далеких от окончательного практического применения, с другой.

В то же время для решения многих практически важных задач машинного зрения общая проблема «понимания изображений» может быть редуцирована к гораздо более простой и ясной - *обнаружения или детектирования* по одному или нескольким изображениям объектов, удовлетворяющих некоторому, заранее известному *модельному* описанию. Теория машинного зрения предлагает целый ряд различных модельных описаний наблюдаемых объектов, которые могут быть использованы для их обнаружения. В литературе описан широкий спектр таких моделей – от простейших *признаковых описаний* до высоко специализированных и изощренных *структурных моделей*. Однако общий метод составления работоспособных моделей по упомянутым причинам отсутствует. Отсюда следует, что разработка и использование мо-

делей, пригодных для эффективного решения задачи обнаружения соответствующих объектов, в значительной степени остается на грани науки и искусства, то есть требует особого "know-how" или, другими словами, знания предметной области, отражающего многолетний опыт исследований по решению частных задач.

Полная автоматизация решения задачи детектирования сложных объектов открывает перед системами «машинного зрения» огромное число потенциальных областей применения, таких как: промышленная инспекция и контроль качества, робототехника, навигация и транспортировка, медицина и биомеханика, инженерный труд, автоматизация проектирования и множество других. При этом задача детектирования является безусловно базовой, но частной технологической задачей по отношению ко всему комплексу основных целевых задач, которые в общих чертах могут быть сформулированы следующим образом:

- обнаружение объектов и изменений в сцене наблюдения;
- высокоточные измерения элементов сцены;
- слежение за объектами;
- самоориентация и самопозиционирование;
- реконструкция поверхностей и обнаружение трехмерных структур;
- описание сцены и идентификация объектов;
- визуализация сцены наблюдения.

Методы машинного зрения

На протяжении 10 – 15-ти последних лет в алгоритмическом аспекте последовательность действий по обработке изображения определяется так называемой *модульной парадигмой*. Эта парадигма, предложенная на основе длительного изучения механизмов зрительного восприятия человека, утверждает, что обработка изображений опирается на несколько последовательных уровней восходящей информационной линии «иконическое представление объектов (растровое изображение, неструктурированная информация) — *символическое представление* (векторные и атрибутивные данные в структурированной форме, реляционные структуры)». Она должна осуществляться по модульному принципу посредством следующих этапов обработки:

- предобработка изображений;
- сегментация;
- выделение геометрической структуры;
- определение относительной структуры и семантики.

Связанные с этими этапами уровни обработки обычно называются обработкой *низкого уровня, среднего уровня, высокого уровня*. В то время как алгоритмы обработки нижнего уровня (фильтрация простых шумов, гистограммная обработка) могут рассматриваться как хорошо проработанные и детально изученные, алгоритмы среднего уровня (сегментация) продолжают сегодня оставаться в центре приложения исследовательских усилий. За последние годы значительный прогресс был достигнут по отношению к проблемам сопоставления точек и фрагментов изображений (*matching*), выделения признаков внутри малых фрагментов, высокой точности 3D-позиционирования точек, что подразумевает соответствующее моделирование и калибровку датчиков и их комбинаций, выделение простых яркостно-геометрических структур типа «точка», «край», «пятно», «прямая линия», «угол».

Методы обработки высокого уровня, относящиеся собственно к «пониманию изображений», находятся еще в начальной фазе развития и по-прежнему представляют собой «вызов» для сообщества исследователей в области компьютерного зрения и искусственного интеллекта. Безусловно, перспектива создания будущих поколений «интеллектуальных машин», в основном, зависит от дальнейшей разработки именно этого круга алгоритмов.

Наиболее интересные с точки зрения конкретных практических приложений методы и алгоритмы детектирования сложных объектов относятся к обработке изображений среднего уровня. Одна из центральных проблем, отличающих методы обработки изображений от хорошо изученной теории обработки сигналов, заключается в разработке методов обнаружения объектов, *слабо чувствительных к разнообразным видам изменчивости*, характерным лишь для изображений, таким как ракурсные и радиометрические искажения, а также различные виды искажений, несводимые к вероятностным моделям (шумы формы). В настоящее время наряду с множеством эвристических алгоритмов обнаружения конкретных типов объектов,

был предложен ряд более общих подходов: методика корреляционного обнаружения, преобразование Hough, морфологические подходы Пытьева и Серра.

Машинное зрение в ряду сопряженных технических дисциплин

Наряду с термином «машинное зрение» часто употребляются такие понятия как: зрение роботов (robot vision), компьютерное зрение (computer vision), обработка изображений (image processing), понимание изображений (image understanding). Оценивая *компьютерное зрение* как научную дисциплину, изучающую теорию и базовые алгоритмы анализа изображений и сцен, *машинное зрение* следует рассматривать сейчас как гораздо более комплексную и технологическую область научных и инженерных знаний, охватывающую все проблемы разработки практических систем: выбор схем освещения исследуемой сцены, выбор характеристик датчиков, их количество и геометрия расположения, вопросы калибровки и ориентирования, выбор или разработка оборудования для оцифровки и процессорной обработки, разработка собственно алгоритмов и их компьютерная реализация и т. д. «Зрение роботов» следует трактовать как более узкую область технологий машинного зрения, а именно ту их часть, которая обеспечивает функционирование систем машинного зрения в условиях жестких временных ограничений. К этому понятию безусловно относятся проблемы разработки основанных на изображениях информационных систем, входящих в состав перспективных систем управления сложными объектами (самолет, автомобиль, системы контроля технологических процессов и др.), так как необходимость формирования обратных связей по результатам обработки входных изображений в системах управления требует их быстрого анализа.

Постоянно встречающийся термин «обработка изображений» в последнее время чаще употребляется не как обозначение научной дисциплины, а как указатель на предметную область. При этом в ряде публикаций наметилась тенденция использования этого термина для обозначения обработки нижнего уровня, когда результатом обработки изображения снова является изображение. В то же время термин «понимание изображений» употребляется для обработки верхнего уровня, часто в контексте применения методов искусственного интеллекта.

Машинное зрение тесно связано с целым рядом смежных дисциплин. Это уже упоминавшиеся ранее *компьютерное зрение*, *искусственный интеллект*, а также *распознавание образов* (pattern recognition) и *цифровая фотограмметрия* (digital photogrammetry).

Цифровая фотограмметрия, бурно развивающаяся последние годы дисциплина, пришла на смену аналитической фотограмметрии. В то время как классическая аналитическая фотограмметрия изучала, в основном, метрические соотношения между точками снимков и реальной сцены, современная цифровая фотограмметрия, особенно так называемая close-range (не очень удачный перевод - короткобазисная фотограмметрия, но другого термина пока нет), ставит самые сложные задачи анализа и 3-D описания сцены по видеоданным оптических сенсоров. Многие ведущие западные ученые в настоящее время используют термины digital close-range photogrammetry и machine vision практически как синонимы. Эта точка зрения получает все большее распространение и в России.

Технологии, средства и практические приложения

В данной статье машинное зрение рассматривалось как отрасль знания и научно-техническая дисциплина. Однако, это отрасль знания, ориентированная, прежде всего, на практическое применение. Действительно, множество зарубежных, а в последнее время и российских фирм, демонстрируют значительный прогресс в решении конкретных практических задач анализа изображений, разработки готовых систем машинного зрения «под ключ», развития конечных технологий и специализированных аппаратных и программных средств.

В этом сборнике представлен ряд разработок фирм «СОБИ Видеоскан» и Институт информационных технологий, которые могут служить примерами успешных практических решений в области машинного зрения, существующих в настоящее время в России.