

О ПРИМЕНИМОСТИ МЕТОДОВ НЕПАРАМЕТРИЧЕСКОГО АНАЛИЗА МНОГОМЕРНЫХ МАССИВОВ ДАННЫХ В СФЕРЕ АВТОМОБИЛЬНОГО ТРАНСПОРТА

Средства искусственного интеллекта обнаруживают широкую область применения: при распознавании текста и речи; медицинской диагностике; спам-фильтрах, в биржевой игре и при управлении собственностью, в обеспечении национальной безопасности. Современные тенденции использования средств искусственного интеллекта (средства распознавания образов, проверки статистических гипотез и обнаружения сигналов, генерация когнитивных образов) для анализа изображений диктуют необходимость программного воплощения этих возможностей.

Весь этот комплекс средств активно внедряется в практику работы прикладных информационных систем управления. Он даже получил специальное название – компьютерное зрение.

Наилучших результатов по выделению и определению объектов контроля можно добиться, анализируя интегрированную информацию об объекте, т.е. используя многомерные массивы данных. Примером многомерных данных является космический снимок географической области, на котором каждому пикселю, определяемому своими координатами, соответствуют, помимо RGB цветов, такие характеристики, как средняя радиояркость в данном диапазоне, коэффициенты спектральной яркости и т.д. Если существуют дополнительные подспутниковые измерения по этим участкам (загазованность местности, концентрация определенных веществ и т.д.), возможно увеличение размерности этого пространства. Таким образом, формируется многомерное пространство данных, где в качестве элемента пространства выступают пиксели или элементарные площадки. Эти участки в данном многомерном пространстве образуют сложные геометрические образования.

Авторами разработана специальная программа непараметрического и когнитивного анализа многомерных данных. Методы непараметрической статистики уже доказали свою широкую применимость, важность и актуальность.

Ранее для создания тонких и глубоких выводов приходилось считать, что наблюдаемые случайные величины имеют предполагаемый вид распределения. Однако, иногда незначительные нарушения могут существенно исказить результаты: привести к смещению оценок, доверительных границ и коэффициентов доверия. Один из способов ослабить эти явления — использование статистических правил.

Классические статистические методы, предназначенные для результатов измерений, хотя и основаны на предположении о нормальности, часто относительно нечувствительны к отклонениям от неё. Непараметрические же методы и упомянутая выше программа были построены специально для того, чтобы обходиться вовсе без этого предположения, причём они широко применимы и просты в работе. Более того, непараметрические методы обладают высокой степенью эффективности, по сравнению с классическими при наличии нормальности, и к тому же часто более эффективны в других ситуациях.

В программе реализованы, для решения различных прикладных проблем, современные методы непараметрической статистики.

Представляется эффективным и целесообразным использование одновыборочной задачи о положении (сдвиге) в качестве основы для решения, в частности, таких задач в области автомобильного транспорта, как:

- анализ рентгеновских, микроскопических изображений материалов и элементов конструкций транспортных средств;
- определение марки автомобиля;
- выделение основных элементов транспортного средства;
- считывание автомобильных номеров;
- выделение дорожного препятствия (по параметрам профильной или несущей проходимости транспортного средства).

Методы решения этих задач предназначены для статистического анализа, направленного в первую очередь на изучение параметра положения - медианы совокупности случайных величин.

Кроме одновыборочных, в программе реализованы двухвыборочные задачи. В двухвыборочной задаче о положении (сдвиге) рассматриваются две случайные выборки: одна — из контрольной совокупности, а другая, независимая от первой, — из рабочей совокупности. На основе этих выборок необходимо выяснить, есть ли эффект изменения, который приводит к сдвигу.

Гипотеза об отсутствии эффекта обработки и есть исходная гипотеза. Она означает, что две выборки можно объединить и рассматривать как единую выборку из одной совокупности.

В двухвыборочной задаче о рассеянии (масштабе) данные представляют собой две независимые случайные выборки, по одной для каждой из двух генеральных совокупностей. На основе выборок необходимо выяснить, есть ли различие в мерах рассеяния (масштаба) этих совокупностей.

Алгоритмы решения двухвыборочных задач могут применяться при анализе двух и более многомерных массивов данных для обнаружения биометрически тождественных и нетождественных объектов.

Эти методы могут найти применение при решении таких задач, как:

- выделение движущихся автотранспортных средств и/или пешеходов (обобщённо — объектов), и слежение за ними;
- распознавание лиц водителя и пассажиров транспортных средств;
- слежение за глазами и направлением взгляда водителя транспортного средства (в частности, автогонщика);
- привязка разномоментных и разномасштабных изображений;
- сшивка маршрутных изображений;

На сегодняшний день интеграция данных об объекте и их последующий анализ в виде многомерных слоёв является принципиально новой задачей. Анализ слоёв многомерных данных не по отдельности, а во взаимозависимости с учётом их когнитивных свойств может выявить новую информацию об объекте — конкретном транспортном средстве, водителе, пассажире.