



из трёх элементов «малое», «среднее» и «высокое», где «малое» это от 0 до 10, «среднее» от 10 до 20, а «высокое» от 20 и больше [3].

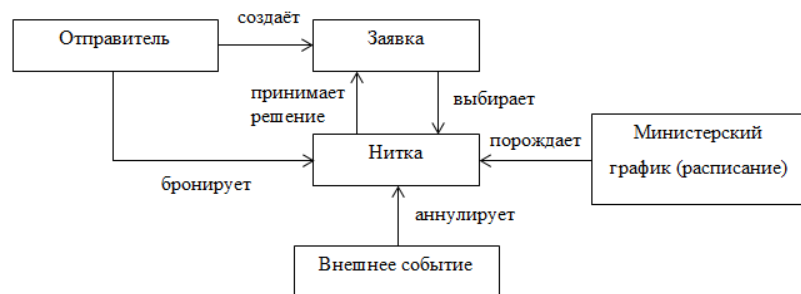


Рис. 1. Диаграмма объектов системы

Помимо количества отказов грузоотправителя в системе учитываются следующие критерии:

- стоимость заявки;
- периодичность заявки;
- количество вагонов;
- наличие опасных веществ;
- габаритность груза;
- дальность перевозки (в сутках);
- наличие скоропортящегося товара;
- количество перевозок грузоотправителя;
- количество отказов грузоотправителя от перевозок.

Возникновения непредвиденного события может повлиять на распределение ниток в системе. Представим ситуацию, в которой для нужд военных были забронированы несколько ниток. В этом случае агенты, соответствующие этим ниткам, должны оповестить все заявки, для которых нитки являются подходящим, что обслуживание их невозможно. После этого агенты «нитка» перестают существовать и учитываться в системе.

3 Структура системы

Система распределения твёрдых ниток графика между грузоотправителями состоит из двух частей: клиентской и серверной.

С помощью *клиента*, устанавливаемого на ПК фирмы, грузоотправитель взаимодействует с заявками, а также бронирует предлагаемые ему нитки графика. Клиентская часть имеет удобный графический интерфейс.

Клиентская часть системы взаимодействует с сервером через Интернет-соединение по протоколу TCP/IP. На *сервере* происходит работа с базами данных, выбор нитки для конкретной заявки и производится оценка эффективности распределения. Структурная схема системы представлена на рисунке 2.

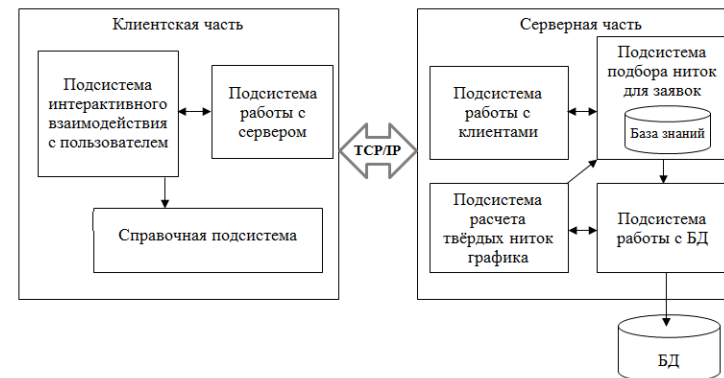


Рис. 2. Структурная схема системы

4 Выводы по работе

Благодаря использованию мультиагентного подхода, разрабатываемая система обладает реактивностью, масштабируемостью и адаптивностью, способна быстро подбирать для грузоотправителя подходящее время в расписании с учетом типов груза, маршрута, количества вагонов и т.д. Внедрение данной системы позволит в разы ускорить процесс бронирования ниток, а также повысить эффективность распределения расписания между грузоотправителями.

Литература

1. Wooldridge, M. Intelligent Agents: Theory and Practice. The Knowledge Engineering Review. – 1995. – Vol.10, №2. – P.115-152.
2. Додонова, Н.Л. Процедура нечеткого вывода состава региональных управляемых промышленных кластеров [Текст] / Н.Л. Додонова, Е.А. Кукольников. – Тольяти: ПВГУС, 2014. – С.52-57.
3. Борисов, В.В. Нечеткие модели и сети [Текст] / В.В. Борисов, В.В. Круглов, А.С. Федулов. – М.: Горячая линия-Телеком, 2007. – 284 с.

Е.Д. Карпова

ВЫБОР АЛГОРИТМА ДЕТЕКТИРОВАНИЯ ОБЪЕКТОВ НА ИЗОБРАЖЕНИИ ДЛЯ РЕШЕНИЯ ЗАДАЧИ ПОИСКА ТУРИСТИЧЕСКИХ ОБЪЕКТОВ НА ФОТОГРАФИЯХ

(Самарский национальный исследовательский университет имени академика С.П. Королёва)

Туризм является важной составной частью экономической системы России. Одним из главных элементов обеспечения комфортного изучения достопримечательностей и различных туристических объектов в городах России, на



данный момент, является карта. Однако по картам нельзя сразу получить интересную информацию об искомом объекте и другие его изображения, также зачастую достопримечательности неузнаваемы для человека, не знакомого с местностью. Поэтому в настоящее время задача детектирования объектов на изображениях является актуальной для распознавания туристических объектов и формирования системы дополненной реальности. Существует множество алгоритмов и подходов для решения задачи детектирования объектов на изображениях, но учитывая специфику конкретной поставленной задачи необходимо рассмотреть лишь некоторые из них, наиболее подходящие для специфики задачи. Особенностью распознавания туристических объектов является то, что эти объекты разнородны (памятники, памятные таблички, здания и так далее) и не поддаются классифицированию.

Алгоритмы для детектирования объектов на изображении

В статье [1] для детектирования объекта на изображении выделяются границы объекта с помощью одного из самых популярных методов – детектора Кенни, алгоритм состоит из пяти отдельных этапов:

1. Сглаживание. Размытие изображения для удаления шума.
2. Поиск градиентов. Границы отмечаются там, где градиент изображения приобретает максимальное значение.
3. Подавление немаксимумов. Только локальные максимумы отмечаются как границы.
4. Двойная пороговая фильтрация. Потенциальные границы определяются порогом
5. Трассировка области неоднозначности. Итоговые границы определяются путём подавления всех краёв, несвязанных с определенными (сильными) границами.

Далее на двоичном изображении, содержащем границы, детектируются структурные элементы (в данном случае окружности) с помощью преобразования Хафа, которое служит для поиска на изображении фигур, заданных аналитически [2]. Исходный алгоритм преобразования Хафа был разработан для поиска прямых линий. Если прямая задана уравнением $y = kx + b$ с параметрами k и b , то пространство параметров будет двумерным. Чтобы использовать преобразование Хафа, исходное изображение контуров бинаризуется и для каждой превосходящей порог точки исходного пространства (x, y) определяется множество прямых, которые могут через нее проходить. Это множество задает в пространстве Хафа некоторую кривую. Данная процедура называется голосованием – каждая точка (x, y) голосует за множество прямых, которые могли бы через нее проходить. Перебор всех точек исходного изображения заполняет аккумулятор, в котором далее ищутся локальные максимумы, которые определяют параметры искомого прямого. Таким образом выделяется объект, структурными элементами которого являются окружности. Очевидным достоинством данного метода является отсутствие необходимости составления обучающей выборки.



В работе [3] для определения марки автомобиля по видеокадру был выбран метод Виолы-Джонса [4], т.к. он обеспечивает необходимую скорость и качество обработки изображений. При обучении использовался метод AdaBoost в варианте, представленном Виолой и Джонсом. Данный алгоритм предназначен для обнаружения объектов определённых классов, изначально для распознавания лиц. Эмблемы автомобилей различных марок отличаются друг от друга, таким образом, их нельзя объединить в один класс объектов и обнаружить их за один проход. Поэтому в итоге формируются отдельные классы для каждого типа эмблем. Плюсом данного подхода является быстрота и качество обработки, но при этом затраты времени на обучение системы увеличивается в разы за счёт увеличения класса объектов, так как для каждого класса необходима своя обучающая выборка.

Для поиска объекта на изображении, метод, описанный в статье [5], использует яркостные признаки. Чтобы устранить влияния освещения и некоторых других особенностей при съёмке, необходимо использовать относительные показатели яркости. Данный метод использует шаблон, который формируется по обучающей выборке. Для описания изображения строится матрица, в которой отображены показатели сравнений областей изображения каждый с каждым. Элементы этой матрицы далее называют признаками. Показатель сходства с шаблоном определяется как отношение совпавших признаков к общему количеству признаков. Недостатком данной методики является чувствительность шаблона к совпадению размерности матриц шаблона и изображения, а также к масштабу и повороту объекта, поэтому необходимо использовать методы предварительной обработки для нормализации геометрии объекта, что влияет на скорость работы алгоритма. Плюсом же является относительная простота этого алгоритма для понимания относительно многих других алгоритмов распознавания объектов на изображении.

Задача заключается в выделении области на фотографии, в которой запечатлён один из известных туристических объектов. Объекты заранее известны, для них могут быть собраны эталонные изображения и подготовлены данные для детектирования объекта. Для различных классов объектов могут применяться разные алгоритмы детектирования с учётом особенностей самих объектов: цвета, текстуры, размеров. Задача может быть упрощена, если известны координаты места, с которого сделан снимок и направление, в котором велась съёмка. Это сводит задачу к перебору достаточно небольшого числа достопримечательностей, поскольку мы можем ограничить поиск объектов некоторым радиусом видимости, который может быть задан константой.

Исходя из всех достоинств и недостатков вышеописанных алгоритмов и методик поиска объекта на изображении, можно сделать вывод, что для решения задачи поиска туристических объектов на фотографии каждый из алгоритмов подходит для решения задачи для некоторых классов объектов, таких как памятные таблички, памятники или здания. Детектор Кенни может быть применён для выделения контуров и последующего их сравнения с шаблонными контурами памятников, так как именно памятники и различные монументы



имеют наиболее уникальные границы. Преобразования Хафа могут быть применены для поиска прямоугольных или круглых областей, с последующим поиском среди них памятных табличек, так как именно они, в основном, имеют заранее чётко известную и простую форму. Методы Виолы-Джонса и относительной яркости подходят для решения задачи в целом, то есть для поиска наиболее распространённых туристических объектов (памятников, монументов, зданий), но для каждого из методов необходимо составить обучающие выборки и обучать систему, кроме того каждый из методов имеет уникальный формат обучающих данных (шаблонов, матриц, коэффициентов обучения). Мы можем сделать вывод о том, что наибольшее количество преимуществ имеет метод Виолы-Джонса, дополнительным его преимуществом является наличие стандартной реализации в библиотеке OpenCV [6], что значительно облегчает его апробацию.

Литература

1. Азаренко, Д.С. Детектирование объекта на изображении и определение его смещения на двух различных изображениях [Текст] /Д.С. Петров //Искусственный интеллект. – 2013. – № 3. – С. 90–97.
2. Корнилов, Ф.А. Поиск прямоугольных контуров на изображениях с помощью преобразования Хафа [Текст] /Ф.А. Корнилов //Современные проблемы математики и её приложений. – 2014. – С. 192.
3. Копылов, И.Е. Определение марки автомобиля по видеок cadру с использованием модифицированного алгоритма Виолы-Джонса [Текст] / И.Е. Копылов, К.Л. Тассов //Наука и образование: электронное научно-техническое издание. – 2012. – №. 06. – С. 205–210.
4. Viola, P. Rapid object detection using a boosted cascade of simple features / P. Viola, M. Jones // Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition. December 2001. Vol. 1. PP. 1063–6919.
5. Агарков, А.В. Поиск объектов на изображениях с помощью относительной яркости отдельных областей [Текст] / А.В. Агарков, Р.О. Кашин, К.С. Ивашко //Искусственный интеллект. – 2014. – № 4. – С. 180–185.
6. Bradski, G. Learning OpenCV: Computer vision with the OpenCV library / G. Bradski, A. Kaehler – O'Reilly. – 2008. – 580 с.

В.П. Клюев

АВТОМАТИЧЕСКАЯ ОБРАБОТКА БОЛЬШИХ ОБЪЕМОВ ДАННЫХ С ИСПОЛЬЗОВАНИЕМ ДЕРЕВЬЕВ РЕШЕНИЙ

(Самарский национальный исследовательский университет
имени академика С.П. Королёва)

В современном мире быстро развиваются технологии автоматического анализа больших объёмов данных. Это вызвано тем, что накопленную инфор-



мацию уже невозможно обработать вручную. Один из методов автоматического анализа данных – это деревья решений.

Построение дерева решений начинается с определения обучающей выборки. Пусть задано некоторое множество X , содержащее примеры, каждый из которых характеризуется n атрибутами. Каждый объект принадлежит некоторому классу C_k .

Возможны 3 ситуации:

1. Множество X состоит из объектов принадлежащих одному классу C_k . Тогда X – это лист, определяющий класс C_k ;
2. Множество X пустое. Тогда X – это снова лист, который принадлежит такому же классу, что и родитель.
3. Множество X состоит из объектов, относящихся к разным классам. Множество X следует разбить на подмножества. Для этого выбирается один из атрибутов, имеющий различные значения $O_1, O_2 \dots O_m$. Множество X разбивается на подмножества $X_1, X_2 \dots X_m$, где каждое подмножество X_i содержит все объекты, имеющие значение O_i для выбранного атрибута. Это процедура повторяется рекурсивно.

После этого возникает несколько вопросов.

Первый вопрос – это выбор атрибута, по которому нужно будет проводить разбиение множества. Это особенно актуально, потому что большинство из известных алгоритмов являются "жадными алгоритмами". Если один раз был выбран атрибут, и по нему было произведено разбиение на подмножества, то алгоритм не может вернуться назад и выбрать другой атрибут, который дал бы лучшее разбиение. Из этого видно, что на этапе построения нельзя сказать даст ли выбранный атрибут оптимальное разбиение или нет.

Второй вопрос – это проблема отсечения избыточных ветвей. Сам по себе вопрос тривиален. Если разбиение множества в узле ведет к листам одного класса, то этот узел можно считать листом этого класса. Но отсюда вытекает другой вопрос.

Третий вопрос – критерий прекращения. Если во множестве на 99 объектов класса C_1 приходится 1 объект класса C_2 , то можно считать этот узел листом класса C_1 ? А если на 95 объектов класса C_1 – 5 объектов класса C_2 ? Другими словами, как определить является узел листом дерева или его нужно еще раз разбить на подмножества.

В настоящее время существует большое число алгоритмов реализующих деревья решений: CART, C4.5, NewId, IRule, CHAID, CN2 и другие. Самыми известными и популярными на сегодняшний день являются алгоритм CART и C4.5.

CART (Classification and Regression Tree) – это алгоритм построения бинарного дерева решений (каждый узел дерева имеет двух потомков). Алгоритм CART использует, так называемый, индекс Gini, который оценивает "расстояние" между распределениями классов.