

УДК 004.896

ИНТЕЛЛЕКТУАЛЬНАЯ ОБРАБОТКА ИНФОРМАЦИИ В СИСТЕМЕ ТЕХНИЧЕСКОГО ЗРЕНИЯ

М.О. Корлякова (*mkorlyakova@yandex.ru*)

П.К. Новиков (*bobersuperzy_b@mail.ru*)

Е.Ю. Прокопов (*evgen-prokopov@rambler.ru*)

А.Ю. Пилипенко (*gamerjd@mail.ru*)

Московский Государственный Технический
университет, Калужский филиал, Калуга

В работе представлены подходы к применению интеллектуальных технологий обработки информации при создании систем технического зрения для бортовых навигационных комплексов. Рассмотрено решение основных задачи обработки изображений и стерео реконструкции с использованием нейросетевой аппроксимации и классификации. Приведены примеры реализации отдельных этапов обработки информации в нейросетевом базисе для бортовых систем технического зрения.

Введение

Автономные мобильные системы все шире включаются в решение практических задач перемещения грузов и патрулирования территорий в естественных средах, которые не подготовлены специально для работы роботов. Широко известны проекты беспилотных автомобилей корпорации Google, конкурс автоботов и роботов, проводимый DARPA[Stavens et al.], система ASIMO(Honda) и другие проекты в области робототехники. Общим для них является направленность на работу на обычных городских или природных площадках, где необходимо определять свое положение и его изменение. Это заставляет использовать для бортовых навигационных систем объекты и детали местности. В качестве средств формирования модели сцены принято использовать весь комплекс доступных датчиков: эхолокацию, лазерные дальнометры, системы технического зрения (СТЗ) и т.п. Однако, в целом ряде случаев применение активных систем наблюдения должно быть ограничено. Например, при решении задач работы в агрессивных средах, где любое внешнее воздействие может спровоцировать ухудшение ситуации или проинформировать противника о проникновении на его территорию.

Тогда единственным способом формирования модели среды становится система технического зрения, которая пассивно «рассматривает» сцену.

Работа СТЗ является вычислительно сложной задачей, требующей значительных ресурсов, которые должны быть предоставлены вычислительной системой мобильного объекта. Кроме того, точность определения положения, передвигающегося в естественной среде транспортного средства, зависит от размера кадра получаемого с камер СТЗ. Незначительное увеличение размера кадра приведет к существенному росту времени обработки, а небольшое уменьшение размера кадра быстро снижает точность. Необходимо обеспечить баланс точности и скорости обработки, поскольку бортовые вычислительные системы подвижных объектов имеют жесткое ограничение ресурсов, но требуют максимально возможной точности. Таким образом, для решения задачи ориентации и навигации с использованием СТЗ на борту автономного мобильного платформы, необходимо искать пути снижения вычислительной сложности задач обработки и анализа изображений.

Применение нейросетевых технологий обработки информации на борту движущегося транспортного средства или в системах навигации роботов позволит повысить производительность бортовой СТЗ, не снижая ее точности. Построение нейросетевой системы технического зрения для решения задачи навигации и ориентации в естественной среде рассмотрим на примере анализа скорости и перемещения мобильного платформы.

Бортовая система технического зрения

Основной подход к решению задачи анализа скорости и перемещения состоит в исследовании снимков сцены, получаемых с нескольких камер СТЗ в процессе движения мобильного платформы. В качестве платформы будем использовать автономное наземное транспортное средство (автомобиль, поезд, мобильный робот и т.п.). Скорость движения платформы не превышает 30 км/ч. В самом простом случае, используем стереопару камер и рассматриваем левый и правый кадры, чтобы по проекциям объектов построить их истинное положение. Камеры СТЗ связаны между собой и откалиброваны. Скорость съемки, как правило, значительно превышает темп обработки информации в СТЗ. Поэтому будем пропускать часть снятых кадров, пока идет обработка последней пары для поиска перемещения и скорости. Общая схема обработки имеет следующий вид:

- получение пары кадров в момент времени t_i и t_{i+1} на борту мобильного объекта;
- выделение характерных объектов сцены и их особых точек (углы, линии, эллипсы) на изображениях в моменты времени t_i и t_{i+1} ,

- т.е. нахождение одинаковых объектов, ракурс которых изменился от одного кадра к другому в бортовой системе зрения;
- определение параметров размещения СТЗ относительно объектов сцены в моменты времени t_i и t_{i+1} , т.е. решение задачи определения трехмерных координат объекта по его проекционной модели – стерео реконструкция;
 - вычисление перемещения платформы на основании информации о координатах СТЗ в пространстве;
 - определение скорости относительно моментов времени t_i и t_{i+1} .

Для решения задачи поиска перемещения в идеальном случае достаточно знания пространственных координат 4-х точек сцены в исходном положении и в текущем. Однако, точное определение координат объектов невозможно, что значительно усложняет поставленную задачу [Жокарева и др., 2012]. Этап калибровки СТЗ реализуется до старта мобильной платформы, а, следовательно, не влияет на скорость решения задачи на маршруте. Другой вопрос, что калибровка неидеальна, а ошибка калибровки включается в общую ошибку определения координат. Помимо этого, ошибка определения истинной координаты складывается из влияния размытия точек (функция размытия точки, зависит от аппаратной части СТЗ), ошибки алгоритма поиска особых точек, которая неизбежно возникает на этапе выделения объектов, ошибки стерео реконструкции и сопоставления координат объектов сцены на этапе вычисления перемещения. Процедуры обработки изображений и стерео реконструкции существенно отличаются с точки зрения формальной математической постановки задачи. Однако они могут быть реализованы за счет обучения нейронных сетей, что позволит строить алгоритм общего решения в однородной форме.

Преобразование и предобработка изображений в нейросетевых системах является наиболее отработанной частью общей схемы обработки приведенной выше. Существует множество примеров использования ИНС в системах технического зрения для разных прикладных задач как на уровне предобработки изображений [Sartin et al., 2014], так и на уровне классификации объектов [Ozbay et al., 2005, Lari et al., 2011].

В рамках поставленной задачи, на этапе предобработки, с точки зрения жесткого ограничения ресурсов и времени для бортовой навигационной системы, наиболее сложным является этап поиска особых точек. Рассмотрим подходы к его нейросетевой реализации в форме сети Хопфилда, которая обеспечит распознавание в условиях зашумленных образов, и сети радиальных нейронов (RBF сети), которая основана на поиске ближайшего шаблона.

Поиск особых точек изображения

В качестве входных данных использовали базу примеров углов и прямых краев, описанных в форме бинарной матрицы. Исходное изображение подвергается контурному анализу с применением фильтра Собеля [Гонсалес и др., 2005], а поиск особых точек осуществляется только для областей, прилегающих к контурам. Пример шаблонов для поиска особых точек приведен на рис. 1.

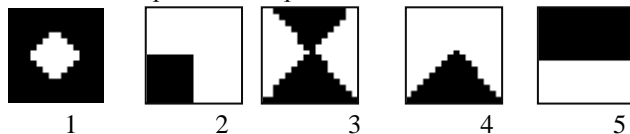


Рис. 1. Примеры шаблонов для поиска углов (1-4) и краев(5).

В процессе распознавания точку кандидата с ее окрестностью размером 16x16 пикселей подают на вход нейросети и оценивают результат на выходе. Эксперименты показали, что обучение сети Хопфилда позволяет определить незначительное количество угловых точек и приводит к значительному (в 3-4 раза по сравнению с RBF сетью) увеличению времени обработки кадров. Сеть радиальных нейронов, сформирована за счет принудительного размещения нейронов в точках соответствующих шаблонам. Она определяет ближайший шаблон и возвращает вектор кода типа образца. Кроме того, число угловых областей выделяемых таким способом значительно больше, чем у сетей Хопфилда. Таким образом, в результате экспериментальной проверки, в качестве метода выделения особых точек принята модель сравнения с шаблонами в сети радиальных нейронов. Пример выбора особых точек (белый «крест») из точек кандидатов («крест») для левого кадра стереопары приведен на рис. 2. а.

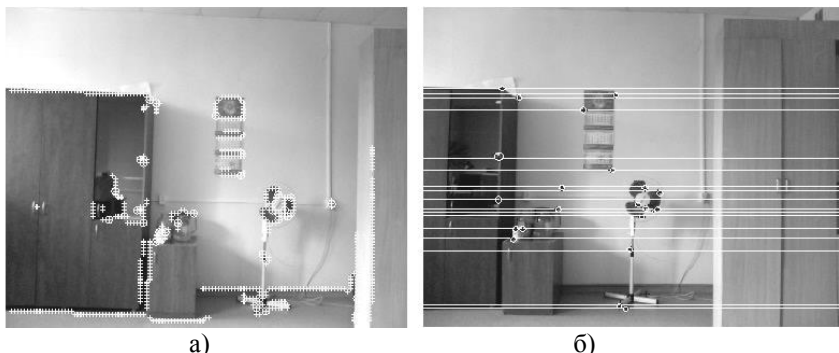


Рис. 2. Поиск особых точек на левом кадре а) и перенос их на правый кадр б).

Каждая особая точка имеет описание растровой области $I_{16 \times 16}$ пикселей, описание градиента изображения $R_{16 \times 16}$ и тип шаблона, который для нее был определен. Это позволяет значительно сократить этап сопоставления особых точек левого и правого кадров, т.к. сопряженные особые точки должны иметь близкий тип, похожие описания областей $R_{16 \times 16}$ и $I_{16 \times 16}$ и лежать в одной эпиполярной плоскости [Кокарева и др., 2012]. Следует отметить, что главным достоинством разработанного алгоритма является его быстроедействие, по сравнению с методом перебора и анализа дескрипторов, но нет инвариантности к поворотам и масштабам.

Стереореконструкция для особых точек изображения

Следующий сложный этап решение задачи стереореконструкции, где из-за ошибок определения положения сопряженных особых точек левого и правого кадров не существует точного решения. В [Кокарева и др., 2012] предложен механизм нейросетевой аппроксимации трехмерных координат точки по ее двум стереопроекциям. Обучение нейросети провели в виртуальной среде с использованием модели СТЗ полученной на этапе калибровки. Для аппроксимации координат использовали сеть типа многослойный персептрон с нелинейным скрытым слоем и 3-мя линейными выходными нейронами, которую обучали на 10000 примерах. Входы нейронной сети: координаты проекций объекта, а выходы его трехмерные координаты. Все примеры были сформированы в виртуальной модели сцены. Эффект достигается за счет переноса основной вычислительной сложности системы на этап обучения нейронной сети, который занимает довольно значительное время (3-5 часов с учетом построения выборки примеров). Кроме того, формальная математическая модель стереореконструкции позволяет работать в любом диапазоне дальностей и ошибки координат будут пропорциональны дальности. Нейросетевой подход требует ограничения диапазона глубины сцены при обучении и решении задачи и показывает резкий рост ошибок за границами этого диапазона. Результаты экспериментов по вычислению координат для виртуальной модели и реальной сцены, показывают, что ошибка стереореконструкции составляет 1-2% дальности объекта. Скорость получения ответа от нейросети $3 \cdot 10^{-4}$ с. и значительно превосходит методы поиска оптимальных координат (0.02 сек.). результаты стереореконструкции сцены с рис.2 приведены на рис.3.а.

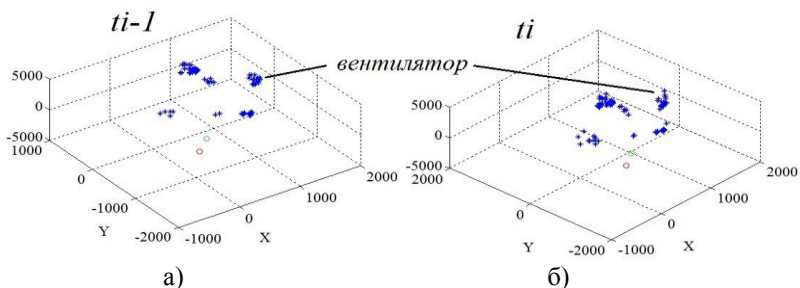


Рис. 3. Стереореконструкция особых точек сцены для момента t_i а) и особых точек сцены для момента t_{i+1} б).

На этапе подготовки стерео реконструкции особые точки сцены отсеиваются на основе эвристического алгоритма для получения наиболее достоверного положения объектов сцены.

Поиск перемещения бортовой системы технического зрения

Стерео реконструкцию особых точек проводим с кадрами стереопары для момента t_i и t_{i+1} (рис.3.а и б). Результаты стерео реконструкции для момента t_i и t_{i+1} образуют множество координат V (положение точек наблюдаемой сцены в момент времени t_i) и V_{new} (положение тех же точек сцены в момент времени t_{i+1}). Перемещение объектов сцены относительно камеры или камеры относительно объектов сцены можно описать уравнением вида:

$$\begin{aligned} M \cdot V &= V_{new} , \\ M &= V_{new} \cdot V^{-1} \end{aligned}$$

где M - матрица переноса-поворота мобильного объекта с СТЗ на борту.

Идеального решения, в силу существования ошибок при определении V и V_{new} , не существует. Построим нейросетевую аппроксимацию этого решения, а поскольку нейронные сети устойчивы к шуму, то возможно получить качественное решение задачи поиска параметров M при ошибках определения координат 1-2% от дальности. В результате экспериментов было показано, что наиболее качественное решение получается за счет разделения общего перемещения на поворот и перенос в следующем виде: $M = R \cdot T$, R – поворот, T – перенос. Причем поиск параметров поворота осуществляем не по V и V_{new} , а по трем ортам направлений наибольшей дисперсии точек сцены для их центрированных

представлений. Тогда частные ошибки отдельных точек из множеств V и V_{new} значительно меньше влияют на общий результат.

Для поиска параметров R использовали многослойный перцептрон с одним нелинейным скрытым слоем. Орты множеств V и V_{new} строим за счет использования метода главных компонент [Ивашина и др., 2012]. Входы нейронной сети содержат координаты ортов центрированного множества точек V_{new} , размещенные в системе координат, заданной ортами множества V . Выходы сети - параметры матрицы поворота R . Все примеры формируем в виртуальной сцене в виде идеального значения R , сопоставленного идеальному значению ортов. К сожалению, устойчивое состояние ортов сцены можно гарантировать только при значительных (несколько тысяч точек) объемах V и V_{new} . На рис.4 приведены ошибки поворота от числа доступных точек для ошибки стерео реконструкции 20 мм, на средней глубине 2 м.

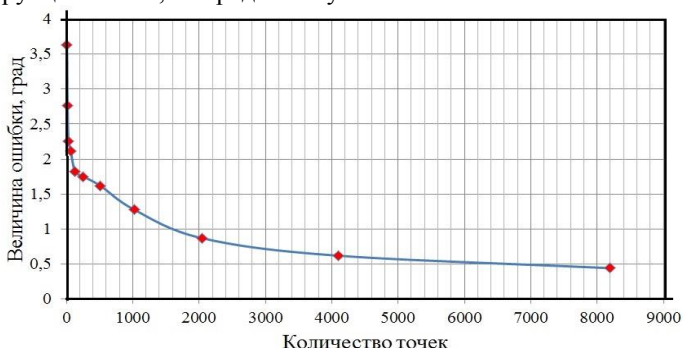


Рис. 2. Ошибка определения поворота от числа доступных точек объектов сцены.

Ошибка поворота становится менее 1 гр. при практически недоступных объемах данных, но 100-200 точек сцены позволяют получать не более чем 1.5-2 гр. ошибки при общем повороте до 90 гр.

Перенос T при известных поворотах определяется за счет анализа V_m, V_{m_new} средних для множеств V и V_{new} , как $T = V_m \cdot R - V_{m_new}$.

Причем, поскольку T зависит от средних по V и V_{new} , то влияние их индивидуальных ошибок практически исключается, а общая ошибка зависит в основном от ошибки определения поворота и составляет 30-40 мм при глубине сцены 2 м. Время обработки первой стереопары, при размере каждого изображения 1280x1024 пикселей, составляет 42 с., а каждой последующей 30 с. при моделировании в среде MatLab. Для

решения аналогичной задачи без использования нейронных сетей пришлось затратить 120-200 с. при эквивалентном уровне ошибок.

Заключение

Результаты моделирования нейросетевой СТЗ показали принципиальную эффективность подхода для решения задач навигации и ориентации в бортовых системах автономных мобильных объектов. Получены вычислительно простые решения задач стерео реконструкции и поиска параметров перемещения. Ошибки традиционной и интеллектуальной модели СТЗ сопоставимы по уровням, а время обработки существенно меньше. Следует отметить, что области эффективной работы нейросетевой модели ограничены, что потребует определения типа среды, где перемещается мобильная платформа.

Однородность решений позволит построить более простую аппаратную и программную реализацию модели в бортовой вычислительной системе робота или автономного транспортного средства.

Список литературы

- [Гонсалес и др., 2005] Вудс Р., Гонсалес Р. Цифровая обработка изображений. – М.: Техносфера, 2005. – 1072с.
- [Ивашина и др., 2012] Ивашина Е.А., Корлякова М.О. Пилипенко А.Ю. Нейросетевое восстановление координат сцены в стереоскопической системе технического зрения// Гибридные и синергетические интеллектуальные системы: теория и практика: материалы 1-го международного симпозиума/ под ред. Проф. А.В. Колесникова. - Калининград: Из-во БФУ им.И.Канта, 2012 — с. 163 -173
- [Кокарева и др., 2012] Кокарева Е.А., Корлякова М.О. Пилипенко А.Ю. Решение задачи стереорекострукции в нейросетевом базисе // XIV Всероссийская научно-техническая конференция «Нейроинформатика-2012»: Сборник научных трудов. В 3-х частях. Ч.1.- М: НИЯУ МИФИ, 2012 - стр. 160-169
- [Lari et al., 2011] Lari Z. and Ebadi H., Automated Building Extraction from High-Resolution Satellite Imagery using Spectral and Structural Information Based on Artificial Neural Networks// International conference on object extraction, vol. 21, no.21. pp.23-27
- [Ozbay et al., 2005] S. Ozbay and E. Ercelebi, Automatic Vehicle Identification by Plate Recognition A Neural Network// World Academy of Science, Engineering and Technology, vol. 9, (2005), pp.222-225
- [Sartin et al, 2014] Sartin Maicon A. and da Silva Alexandre C. R., Evaluation of Image Segmentation and Filtering With Ann in the Papaya Leaf// International Journal of Computer Science & Information Technology (IJCSIT), Vol. 6, February 2014, pp.47-58
- [Stavens et al.] Stavens D. and Thrun S. A Self-Supervised Terrain Roughness Estimator for Off-Road Autonomous Driving // <http://arxiv.org/1206.6872>