Нейросетевой поиск особых точек для системы технического зрения при определении перемещения мобильной платформы

© М.О. Корлякова, Е.Ю. Прокопов, П.К. Новиков

КФ МГТУ им. Н.Э. Баумана, Калуга, 248000, Россия

Представлена разработка модели технического зрения образов для решения задачи навигации мобильного робота в естественной среде. Рассмотрены возможности использования различных способов поиска особых точек. Определены требования к бортовым системам поиска особых точек. В качестве основного подхода предложено использовать нейросетевые модели обработки информации, показаны преимущества и недостатки этого подхода. Проведено исследование возможностей сетей Хопфилда и сетей на основе радиальных нейронов. Проведено моделирование и определены наиболее эффективные по точности и скорости модели сетей. Полученная модель проверена на реальных примерах. Рассмотрены возможные методы ускорения процедуры поиска особых точек вследствие сокращения исследуемых областей. Определены пути реализации предлагаемого подхода в рамках бортовых систем технического зрения

Ключевые слова: нейронные сети, распознавание образов, компьютерное зрение.

В настоящее время компьютерное зрение — одно из самых актуальных направлений развития интеллектуальных цифровых компьютерных технологий. Главная отличительная черта систем технического зрения (СТЗ) — извлечение полезной для нас информации из изображения или последовательности изображений. Компьютерное зрение сосредоточивается на обработке трехмерных сцен, спроецированных на одно или несколько изображений. По одному или нескольким изображениям можно восстановить структуру или другую информацию о трехмерной сцене. Системы технического зрения применяют для построения трехмерных компьютерных моделей зданий и сооружений, для интерпретации фотографий, сделанных со спутника, при работе с медицинскими изображениями, для отслеживания положения мобильных роботов и т. д. [1]. Рассмотрим задачу навигации мобильного робота в естественной среде на основе использования видеопотока от пары связанных камер. Требуется извлечь информацию о положении и перемещении объектов или мобильного робота из информации об изменении состояния стереопары кадров. Исходя из характера задачи, необходимо учесть следующие ограничения:

информация о характеристиках фона и наблюдаемых объектах чаще всего содержит неточные размеры объектов или вообще отсутствует;

в связи с ограниченностью времени на принятие решения, анализ и распознавание должны выполняться с минимальными задержками по времени;

работа системы слежения и обработки должна проводиться при минимальном участии человека или в автономном режиме.

Общую схему решения задачи определения перемещения мобильного робота в среде можно разделить на следующие этапы:

1) выделение характерных объектов сцены на изображениях в моменты времени t_i и t_{i+1} , т. е. нахождение одинаковых объектов, ракурс которых изменился от одного кадра к другому;

2) определение координат особых точек мишени или объектов сцены на проекциях, генерируемых камерам СТЗ;

3) определение параметров размещения СТЗ относительно мишени в моменты времени t_i и t_{i+1} , т. е. решение задачи нахождения трехмерных координат объекта по его проекционной модели;

4) вычисление перемещения платформы на основании информации о координатах СТЗ в пространстве в соответствии с матричным уравнением вида

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

где $\mathbf{M}$ — матрица переноса-поворота объекта сцены (или обратная матрица передвижения системы координат, связанных с СТЗ); V — положение точек наблюдаемой сцены в момент времени t_i ; V_{new} — положение тех же точек сцены в момент времени t_{i+1} . Матрица $\mathbf{M}$ содержит 12 неизвестных параметров, однозначно определяющих изменение положения объекта по углу поворота и переносу в глобальной системе координат;

5) определение угловой и линейной скорости относительно моментов времени t_i и t_{i+1} и параметров матрицы $\mathbf{M}$.

Таким образом, для решения задачи в идеальном случае достаточно знания пространственных координат четырех точек сцены в исходном и текущем положениях. Однако, как показано в работе [2], точное определение координат объектов невозможно, что значительно усложняет поставленную задачу. Для повышения качества результатов принято использовать значительно большее множество точек сцены, что приводит к увеличению времени на решение общей задачи. Причем следует отметить, что этапы 1 и 2 имеют наибольшую вычислительную сложность. Таким образом, необходимо разработать точную и производительную систему сопоставления кадров стереопары.

Данная система проводит обработку и анализ изображений, поступающих от оптико-электронных систем, в следующем порядке (рис. 1):

Рис. 1. Общая схема обработки стереопары кадров

1) захват и предварительная обработка изображений, которая подразумевает получение очередной пары кадров из стереопары, их предобработку и выделение наиболее интересных областей (кандидатов на принадлежность к особым точкам) из одного кадра стереопары;

2) определение парных точек изображений стереопары (т. е. точек одного объекта на левом и правом кадре) для текущей пары кадров в результате поиска сопряженных точек правого кадра относительно выделенных особых точек левого кадра;

3) сопоставление стереопар, получаемых в моменты времени t_i и t_{i+1} (или выделение ключевых точек видеопоследовательности).

Рассмотрим последовательно все этапы решения поставленной задачи.

1. Захват и предварительная обработка изображений. Для захвата изображения использовали две камеры HP Deluxe Webcam, установленные на жестко фиксированном расстоянии. Поскольку система должна работать с минимальными задержками во времени, этот факт накладывает ряд ограничений на размер входного изображения в пикселах. Однако уменьшение размера напрямую связано со снижением качества распознавания. Основываясь на данном факте, было принято решение считать ограничения времени реакции системы менее важным критерием ее качества. Таким образом, входными данными системы являются кадры стереопары — каждое изображение размером 1280×1024 пикселей.

В процессе предварительной обработки все кадры подвергали переводу в полутоновое представление, а над «левым» кадром стереопары проводили следующие дополнительные операции:

сжатие изображения;

определение контура объектов.

Это необходимо для увеличения скорости поиска областей кандидатов, так как скорость обработки черно-белого одноканального изображения намного выше, чем скорость обработки цветного трехканального изображения.

Определение контура объектов значительно сужает область поиска особых точек, поскольку там, где нет контуров, нет и особых точек изображения. Как правило, это области с монотонным изменением интенсивности, что не позволяет проводить процедуры сопоставления проекционных образов точек сцены.

Положения границ областей интенсивности определяли с помощью оператора Собеля. По сути, это дискретный дифференциальный оператор, вычисляющий приближенное значение градиента яркости изображения. Результатом применения оператора Собеля в каждой точке изображения является вектор градиента яркости в этой точки. Оператор Собеля основан на свертке изображения небольшими сепарабельными целочисленными фильтрами в вертикальном и горизонтальном направлениях, поэтому его относительно легко вычислить. Оператор исполь-

зует ядра 3×3 , с которыми сворачивают исходное изображение для вычисления изображения, где каждая точка содержит приближенные производные по x и по y . Пусть I — исходное изображение, а G_x, G_y — два изображения, где каждая точка содержит приближенные производные по x и по y . Вычисляют их следующим образом:

$$G_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} * I, \quad G_y = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} * I;$$

где «*» обозначает двумерную операцию свертки [3].

Результатом является сжатое изображение стереопары с выделенной границей. Пусть области исходного изображения, покрытые границами, являются кандидатами на принадлежность к особым точкам.

Распознавание изображений объектов наиболее сложный процесс, поскольку подразумевает значительный перебор областей в изображении. Зачастую качество распознавания зависит от эффективности множества процессов, начиная от процесса ввода изображения и заканчивая формированием признаков объекта. Поэтому имеет смысл ограничить области интереса в изображениях большого размера. В качестве областей, используемых для сравнения кадров, были выбраны точки с наибольшим перепадом градиента яркости. Наиболее ярким примером особых точек являются углы на изображении сцены. Их определяли методом сравнения с эталоном. Все граничные точки сопоставляли с эталонами для нахождения максимально похожего.

Поиск угловых точек можно реализовать в рамках различных подходов с учетом размеров найденных областей и их ориентации. Примером может служить детектор углов Хариса [4], который построен на исследовании монотонности областей изображения. Однако, поскольку вся система формируется с ориентацией на нейросетевое решение задачи, для ускорения и упрощения поиска было предложено использовать нейросеть Хопфилда и сети радиальных нейронов. При обучении обеих сетей использовали образцы углов, прямых краев и монотонных областей, повернутых под различными углами. Примеры образцов приведены на рис. 2, где показаны 5 из 55 принятых для обработки эталонов. Каждый эталон представлен бинарной матрицей размером 16×16 пикселей.

Рис. 2. Примеры шаблонов целевых областей

В процессе распознавания точку с ее окрестностью размером 16×16 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 (до 120 с.) увеличению времени обработки кадров. Сеть радиальных нейронов построена на вычислении евклидовой нормы расстояния от области вокруг исследуемой точки до всех эталонов по следующей схеме:

$$d(p, q_j) = \sqrt{\sum_{k=1}^{n=256} (p_k - b_{jk})^2},$$

где $p = \langle a_{1,1}, a_{1,2}, \dots, a_{1,16}, a_{2,1}, \dots, a_{2,16}, \dots, a_{16,16} \rangle$ — векторизованная матрица области вокруг точки; $a_{jk} = [0,1]$ — значение пиксела исследуемой области, приведенное к интервалу $[0,1]$; $q_j = \langle b_{j,1,1}, b_{j,1,2}, \dots, b_{j,1,16}, b_{j,2,1}, \dots, b_{j,2,16}, \dots, b_{j,16,16} \rangle$; $b_{jk} = \{0,1\}$ — значение пиксела эталона.

Если минимальное евклидово расстояние достигается при сравнении с шаблонами, соответствующими угловым особенностям, то точка считается особенной, иначе она отсеивается. После определения набора особых точек, их координаты пересчитываются по отношению к масштабу исходного, не сжатого изображения. Время сопоставления с эталонами в сети радиальных нейронов не превышает 20...30 с, что позволяет говорить о его превосходстве. 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 эталонами в сети радиальных нейронов.

На рис. 3 представлен результат обработки «левого» кадра первой стереопары — поиск ключевых точек. Крестиками здесь отмечены точки, в которых градиент яркости имеет максимальное значение. Эти точки были отфильтрованы простым перебором и сопоставлением с эталоном. В результате образовался набор особых точек, отмеченных на рис. 3 кружками).

Таким образом, результатом этапа 1 является набор координат особых точек для «левого» кадра первой стереопары.

2. Определение сопряженных точек изображений стереопары для текущей пары кадров. Поскольку на предыдущем этапе был получен набор особых точек только для одного из кадров стереопары, то следующим этапом является сопоставление таких точек в пределах одной стереопары. Поиск особых точек изображений для стереопар основан на исследовании пиксельного соответствия отдельных областей кадра [4]. Существующие методы поиска особых точек позволяют найти соответствия при различных ракурсах, масштабах и искажениях. Стереопара,

для которой формируем алгоритм поиска особых точек, имеет практически параллельные оптические оси и идентичные камеры, что позволяет исследовать кадры попиксельно. В рассматриваемом случае для определения сходства, как и ранее, используем евклидово расстояние.

Рис. 3. Выделение особых точек изображения из множества областей кандидатов с использованием сети радиальных нейронов

Пусть $l(x, y)$ — набор особых точек «левого» кадра, а $r(x, y)$ — набор соответствующих точек на «правом» кадре. Сопоставление точек будем проводить на основании сравнения градиентов окружающих точку областей. Для этого рассчитаем градиент в найденной точке $l^*(x, y)$, а затем, перебирая все возможные точки $r_j(x, y)$, найдем такую точку $r^*(x, y)$, для которой евклидово расстояние между градиентами точек $l^*(x, y)$ и $r^*(x, y)$ является минимальным, т. е. $l^*(x, y)$ является отображением $r^*(x, y)$, если

$$\min \sqrt{\left[\sum \left(\frac{\frac{dL_i}{dx}}{\max \frac{dL_i}{dx}} - \frac{\frac{dR_j}{dx}}{\max \frac{dR_j}{dx}} \right) \right]^2 + \left[\sum \left(\frac{\frac{dL_i}{dy}}{\max \frac{dL_i}{dy}} - \frac{\frac{dR_j}{dy}}{\max \frac{dR_j}{dy}} \right) \right]^2} =$$

$$= \sqrt{\left[\sum \left(\frac{\frac{dL_i}{dx} \Big|_{i=l^*}}{\max \frac{dL_i}{dx} \Big|_{i=l^*}} - \frac{\frac{dR_j}{dx} \Big|_{i=r^*}}{\max \frac{dR_j}{dx} \Big|_{i=r^*}} \right) \right]^2 + \left[\sum \left(\frac{\frac{dL_i}{dy} \Big|_{i=l^*}}{\max \frac{dL_i}{dy} \Big|_{i=l^*}} - \frac{\frac{dR_j}{dy} \Big|_{i=r^*}}{\max \frac{dR_j}{dy} \Big|_{i=r^*}} \right) \right]^2},$$

где $i = \overline{1, n}$, $j = \overline{1, m}$.

Однако исследование всех точек «правого» кадра займет значительное время. Для повышения скорости обработки при построении градиентов «правого» кадра следует пропускать монотонные области. Отфильтровать их можно, рассчитав среднее квадратическое отклонение σ между горизонтальным и вертикальным градиентами (если $\sigma < 1$, то область является монотонной). Кроме того, для повышения скорости обработки и фильтрации ложных особенностей будем сравнивать только области вдоль эпиполярных линий [3].

Рассмотрим пример. Пусть имеются две камеры (рис. 4). Точка пространства X проецируется в x на плоскость изображения левой камеры и в x' на плоскость изображения правой камеры. Прообразом точки x на изображении левой камеры является луч xX . Этот луч проецируется на плоскость правой камеры в прямую l' , называемую эпиполярной линией. Образ точки X на плоскости изображения правой камеры обязательно лежит на эпиполярной линии l' .

Рис. 4. Схема определения положения эпиполярной линии стереопары (C, C' — центр первой и второй камеры соответственно)

Таким образом, каждой точке x на изображении левой камеры соответствует эпиполярная линия l' на изображении правой камеры. При этом пара для x может лежать только на соответствующей эпиполярной линии.

На рис. 5 представлен результат работы алгоритма сопоставления точечных особенностей. Здесь выделены эпиполярные линии, а также крестиками отмечены точки, соответствующие особым точкам левого кадра.

Стоит отметить, что на рис. 2 и 4 имеются ложные соответствия. Для их отсека используют алгоритмы фильтрации и расширения (рис. 6). При этом в каждой области соответствия «левого» и «правого» кадров осуществляется поиск точного попиксельного сопоставления элементов изображений, что позволяет значительно увеличить число сопряженных точек и повысить точность определения их координат.

3. Сопоставление стереопар, получаемых в моменты времени t_i и t_{i+1} . Для сопоставления следующей стереопары в качестве опорных точек используются особые точки предыдущей стереопары.

Алгоритм сопоставления также основан на вычислении евклидова расстояния между разностями градиентов. Однако использовать в качестве фильтра эпиполярную линию здесь нельзя, так как за промежутки времени $t_{i+1} - t_i$ (полное время обработки стереопары от

Рис. 5. Размещение эпполярных линий в кадре и поиск сопряженных особых точек стереопары

Рис. 6. Поиск особых точек для кадров с большим разрешением после определения сопряженных областей «левого» и «правого» кадров

момента захвата изображения до получения координат соответствующих точек и следующего захвата) сцена меняется. Несмотря на это, оптимизировать алгоритм можно, если осуществлять поиск новых координат относительно координат предыдущей стереопары в области, размер которой не превышает максимально возможное смещение объектов за время между кадрами.

Экспериментально было установлено, что для сцены, самая дальняя точка которой расположена в 4 м от положения камер, при данной скорости обработки координаты искомой точки лежат в радиусе 10 пикселей. Поэтому, рассматривая окрестность 20×20 пикселей относительно предыдущих координат точки, можно получить координаты точки текущей стереопары.

Таким образом, разработанный алгоритм обеспечивает сопоставление особых точек стереопар с высокой точностью, что позволяет снизить ошибки следующих этапов стереорекострукции. Показано преимущество системы прямого доступа к эталонам на основе сети радиальных нейронов.

Время обработки первой стереопары, размер каждого изображения которой, 1280×1024 пиксела, составило 42 с, а каждой последующей — 30 с.

Все составляющие алгоритма были реализованы на языке C++ и включены в состав программного обеспечения СТЗ. В дальнейшем разработанную систему планируется использовать для определения расстояния до объекта при измерении его положения, скорости и направления движения.

ЛИТЕРАТУРА

- [1] Девятериков Е.А., Михайлов Б.Б. Система технического зрения для измерения пути мобильного робота. *Сб. тр. науч.-техн. конф. «Техническое зрение в системах управления-2011»*. Москва, ИКИ РАН, 2012, с. 219–224.
- [2] Ивашина Е.А., Корлякова М.О., Пилипенко А.Ю. Формирование признаков для нейросетевого поиска направления перемещения объекта в сцене. *XV Всерос. науч.-техн. конф. «Нейроинформатика-2013»*. Сб. науч. тр. Москва, НИЯУ МИФИ, 2013, ч. 2, с. 40–49.
- [3] Кокарева Е.А., Корлякова М.О. Пилипенко А.Ю. Решение задачи стереорекострукции в нейросетевом базисе. *XIV Всерос. науч.-техн. конф. «Нейроинформатика-2012»*. Сб. науч. тр. В 3 ч. Москва, НИЯУ МИФИ, 2012, ч. 1., с. 160–169.
- [4] Шапиро Л., Стокман Дж. *Компьютерное зрение*. Москва, БИНОМ. Лаборатория знаний, 2006, 752 с.

Статья поступила в редакцию 05.06.2014

Ссылку на эту статью просим оформлять следующим образом:

Корлякова М.О., Е.Ю. Прокопов, П.К. Новиков. Нейросетевой поиск особых точек для системы технического зрения при определении перемещения мобильной платформы. *Инженерный журнал: наука и инновации*, 2014, вып. 6. URL: <http://engjournal.ru/catalog/pribor/robot/1274.html>

Корлякова Мария Олеговна родилась в 1969 г., окончила КФ МГТУ им. Н.Э. Баумана в 1992 г. Канд. техн. наук, доцент кафедры «Системы автоматического управления» КФ МГТУ им. Н.Э. Баумана. Область научных интересов: нейронные сети, интеллектуальные системы управления. e-mail: mkorlyakova@yandex.ru

Новиков Павел Константинович родился в 1991 г. Студент кафедры «Системы автоматического управления» КФ МГТУ им. Н.Э. Баумана. Область научных интересов: распознавание образов, нейронные сети. e-mail: evgen-prokopov@rambler.ru

Прокопов Евгений Юрьевич родился в 1992 г. Студент кафедры «Система автоматического управления», КФ МГТУ им. Н.Э. Баумана. Область научных интересов: распознавание образов, нейронные сети. e-mail: evgen-prokopov@rambler.ru

Neuronet search of special points for computer vision system when determining mobile platform movements

© M.O. Korlyakova, E.Yu. Prokopov, P.K. Novikov

¹ Kaluga Branch of Bauman Moscow State Technical University, Kaluga, 248000, Russia

The main purpose of the article is to present a computer vision system model for solving the problem of mobile robot navigation in natural environment. We consider the possibilities of using different ways of searching certain points and set the requirements to onboard systems. As the main approach we offer to use neuronet models of information processing, and show advantages and disadvantages of this approach. The article examines Hopfield networks abilities and networks on the basis of radial neurons. We carried out modeling and determined the most effective networks models in terms of accuracy and speed. The received model is checked on real examples. We studied possible methods of certain points search acceleration and defined the ways of implementing the proposed approach within onboard systems of computer vision system.

Keywords: neuron networks, object recognition, computer vision system

REFERENCES

- [1] Devyaterikov E.A., Mikhailov B.B. Sistema tekhnicheskogo zreniya dlya izmeneniia puti mobil'nogo robota [The vision system for measuring path of the mobile robot]. *Sbornik trudov nauch.-tekhn. konf. «Tekhnicheskoe zrenie v sistemakh upravleniya–2011»* [Proc. of the sci. and eng. conf. "Technical Vision in Control Systems'2011"]. Moscow, SRI RAS, 2012, pp. 219–224.
- [2] Ivashina E.A., Korlyakova M.O., Pilipenko A.Yu. Formirovanie priznakov dlya neirosetevogo poiska napravleniya peremeshcheniya ob"ekta v stsene [Formation of signs for the neural network search of the direction of the object movement in the scene]. *XV Vserossiyskaya nauchno-tekhnicheskaya konferentsiya «Neiroinformatika–2013»: Sbornik nauchnykh trudov* [XV All-Russian Scientific and Technical Conference "Neuroinformatics'2013": Collection of scientific papers]. In 3 parts. Part 2. Moscow, NRNU MEPhI, 2013, pp. 40–49.
- [3] Kokareva E.A., Korlyakova M.O., Pilipenko A.Yu. Reshenie zadachi stereorekonstruktsii v neirosetevom bazise [Solution of stereo reconstruction problem in the neural network basis]. *XIV Vserossiyskaya nauchno-tekhnicheskaya konferentsiya «Neiroinformatika–2012»: Sbornik nauchnykh trudov*. [XIV All-Russian Scientific Conference "Neuroinformatics'2012": Collection of scientific papers]. In 3 parts. Part 1. Moscow, NRNU MEPhI, 2012, pp. 160–169.
- [4] Shapiro L., Stokman J. *Kompyuternoe zrenie* [Computer vision]. Transl. from Engl., Moscow, BINOM, Laboratoriia znanii Publ., 2006, 752 p.

Korlyakova M.O. (b. 1969) graduated from Kaluga Branch of Bauman Moscow State Technical University in 1992. Ph.D., Assoc. Professor of the Automatic Control Systems Department at Kaluga Branch of Bauman Moscow State Technical University. Scientific interests include neural networks, intelligence control systems. e-mail: mkorlyakova@yandex.ru

Novikov P.K. (b. 1991) is a student of the Automatic Control Systems Department at Kaluga Branch of Bauman Moscow State Technical University. Scientific interests include neural networks, pattern recognition. e-mail: evgen-prokopov@rambler.ru

Prokopov E.Yu. (b. 1992) is a student of the Automatic Control Systems Department at Kaluga Branch of Bauman Moscow State Technical University. Scientific interests include neural networks, pattern recognition. e-mail: evgen-prokopov@rambler.ru