

УДК 004.896

## ПОДХОД К НАСТРОЙКЕ СИСТЕМЫ ТЕХНИЧЕСКОГО ЗРЕНИЯ ДЛЯ МОБИЛЬНОЙ ПЛАТФОРМЫ

Е.А. Ивашина (*mkorlyakova@yandex.ru*)  
Калужский филиал МГТУ им.Н.Э.Баумана, Калуга

М.О. Корлякова (*mkorlyakova@yandex.ru*)  
Калужский филиал МГТУ им.Н.Э.Баумана, Калуга

А.Ю. Пилипенко (*mkorlyakova@yandex.ru*)  
Калужский филиал МГТУ им.Н.Э.Баумана, Калуга

А.А. Филимонков (*mkorlyakova@yandex.ru*)  
КБ-683, филиал ФГУП НПЦ АП им.Н.А.Пилюгина – «СПЗ»,  
Калуга

**Аннотация.** В статье рассмотрены возможности интеллектуализации систем управления беспилотными транспортными средствами за счет использования систем технического зрения и определены основные особенности и проблемы этого подхода. Рассмотрен нейросетевой подход к построению интеллектуальной модели формирования параметров алгоритма обработки кадров в процессе движения беспилотного наземного модуля в естественной среде. Сформулированы основные правила обучения и проведено моделирование движения с обучением и без обучения. Показано, что периодическое самообучение нейронной сети, проводимое на борту автономной беспилотной системы в процессе движения, позволяет повысить точность вычисления пройденного пути и координат в пространстве за счет постоянной адаптации к изменяющимся условиям среды.

**Ключевые слова:** нейронные сети, системы технического зрения.

### Введение

Беспилотные транспортные средства позволяют решать задачи наблюдения за территориями, транспортировки грузов и обеспечения безопасности людей. Повышение автономности беспилотных систем основано на построении качественной бортовой системы анализа внешней среды. Причем желательно иметь возможность вычислить свое положение и определить состав объектов внешней среды и их поведение. Например установить, что обнаруженное препятствие изменило свой цвет

(светофор). Следует отметить, что современные навигационные комплексы решают с высокой точностью задачи определения положения объекта и ориентации его основных осей в пространстве. Однако, с течением времени накапливающиеся ошибки снижают качество результатов и обычно могут характеризовать объекты среды только их положением относительно беспилотного транспортного средства. Кроме того, самые распространенные бортовые навигационные системы – бесплатформенные инерциальные навигационные системы (БИНС) показывают растущие во времени ошибки и при нормальном движении платформы. Например, момент, когда беспилотное наземное мобильное устройство стоит или проскальзывает на месте, может быть интерпретирован измерителями как движение. Поэтому, как правило, применяют дополнительные методы коррекции результатов по пути следования. Объединение нескольких независимых источников измерительной информации позволяет существенно повышать качество ориентации и навигации в естественных условиях. Наиболее естественные средства наблюдения среды, с точки зрения человека, связаны с системами технического зрения (СТЗ), которые позволяют интерпретировать цвет, форму и расстояние-размеры объекта на основе анализа видеопотока. Кроме того, использование изображений сцены позволяет существенно повысить интеллектуальные возможности навигационной системы, т.к. позволяет идентифицировать однородные объекты сцены (например, отдельные автомобили в потоке) и предсказывать их поведение.

Проблемой СТЗ является их производительность и точность. В настоящее время использование обучаемых [Кокарева и др., 2012] и адаптируемых СТЗ(SLAM)[Montemerlo et al, 2002] позволяет снизить вычислительную сложность задачи визуального анализа сцены и повысить ее точность на этапе вычисления координат. Однако, работающая в естественных сценах СТЗ, подвергается существенным внешним воздействиям, что приводит к потере непрерывности измерений еще на этапе обработки кадра. В условиях реальных сцен на вид объектов оказывает влияние время суток и сезон, погода, искусственные источники света, состав объектов сцены и тому подобное. Поэтому СТЗ должна отслеживать эти изменения и уметь их запоминать, т.е. накапливать опыт. Таким образом, необходимо сформировать модель интеллектуальной настройки процедур обработки кадров.

## **1 Нейросетевая система технического зрения**

Рассмотрим более подробно особенности решения задачи одометрии в СТЗ. Определение перемещения автономного беспилотного объекта

обычно реализуется в моно или стерео камерной системе технического зрения и связано с обработкой последовательных кадров видеопотока. Примеры таких систем активно разрабатываются и внедряются (поиск препятствий на железнодорожных путях [Иванов, 2011], управление движением мобильного робота [Девятериков и др, 2013] и т.п.). С другой стороны, обучаемые модели принятия решений и распознавания образов позволяют решать сложные задачи обработки информации и распознавания образов за счет обобщения опыта. Примеры таких систем можно видеть в [Ozbay et, 2005, Кокарева и др., 2012]. Таким образом, для повышения интеллектуальности СТЗ и снижения ее вычислительной сложности с сохранением точности будем использовать нейросетевую реализацию СТЗ, общая схема обработки информации в которой выглядит следующим образом:

- получение пар кадров в момент времени  $t_i$  и  $t_{i+1}$  на борту мобильного объекта;
- сжатие кадров к масштабу  $1/16$ ,
- поиск сопряженных пар точек в сжатых кадрах ( $N$  - число найденных пар) на основе последовательности операций:
  - линейное контрастирование,
  - выделение края (параметр  $g$  – уровень отсечения по модулю градиента изображения),
  - выделение угловых особенностей за счет нейросетевого фильтра [Корлякова и др, 2014],
  - сопоставление фрагментов за счет анализа уровня корреляции изображения ( $k$ ) для угловых особенностей по кадрам в момент времени  $t_i$  и  $t_{i+1}$  [Сырямкин и др, 2010].
- восстановление координат найденных пар точек до полного масштаба,
- повторный поиск сопряженных пар точек для окон размера  $32 \times 32$  вокруг каждой точки нижнего уровня поиска,
- определение трехмерных координат для пар точек в момент времени  $t_i$  и  $t_{i+1}$  за счет нейросетевой стерео реконструкции [Кокарева и др., 2012],
- определение перемещения от момента времени  $t_i$  к  $t_{i+1}$  в виде параметров матрицы  $M$  - поворота-переноса СТЗ.

Проблема простой модели обработки в том, что даже незначительная разница в интенсивности фрагментов приводит к резкому снижению числа  $N$  – найденных сопряженных пар точек стереопары. Пример приведен на

рис.1, где рис.1.а соответствует  $N=107$  и фрагменты правых кадров а.2 и а.4 отличаются от левых (а.1 и а3), но близки между собой, а рис.1.б порождает  $N=0$  пар сопряженных точек, где б.4 резко отличен от остальных фрагментов.

Интеллектуализация СТЗ связана с адаптацией фильтров на этапе обработки кадров к текущей сцене [Соколов 2011]. Существенные различия кадров в видеопотоке возникают за счет резких изменений освещенности, появления «быстрых» объектов и т.п. Решение этой проблемы лежит либо в разработке сравнительно сложных фильтров и следующих за ними механизмов сопоставления кадров, либо в получении простого алгоритма обработки с изменяемыми параметрами.

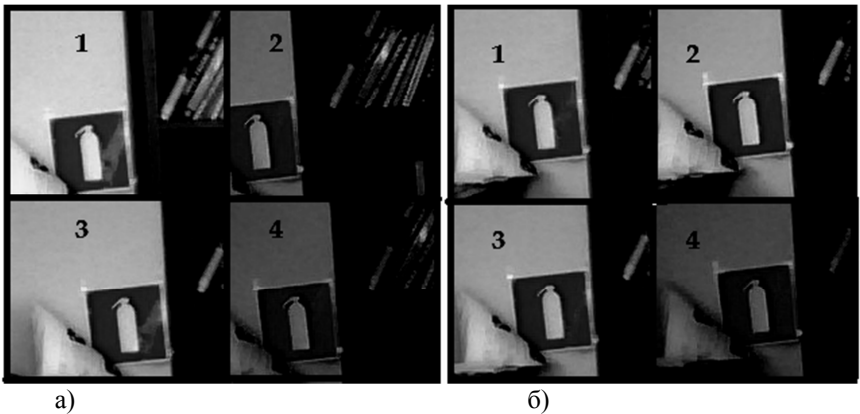


Рис. 1. Результаты поиска пар особых точек в моменты времени  $t_i$  (изображения 1 и 2) и  $t_{i+1}$  (изображения 3 и 4) для кадров а) близкой интенсивности и б) различной интенсивность правых кадров (2 и 4 изображения).

## 2 Формирование обучаемой модели настройки системы технического зрения

Процедура изменения параметров должна обеспечить необходимые значения выходных параметров для результатов обработки последовательных кадров. Для задачи навигации и ориентации беспилотной мобильной платформы основной выходной характеристикой обработки кадров стереопары является число обнаруженных сопряженных точечных особенностей изображения [Корлякова и др, 2014] и их точность. В рамках подхода к стерео реконструкции и поиску параметров перемещения СТЗ представленных в [Кокарева и др., 2012], необходимо

обеспечить не менее 100 пар сопряженных точек.

Для оценки качества обработки и поиска параметров алгоритма обработки кадров предлагается использовать обучаемую модель в составе:

- учитель  $T(Ip(n), n = 1..C, P = \langle g, k, m1, m2, m3, m4 \rangle, A)$ , где  $P$  – вектор параметров алгоритма предобработки (в текущей модели  $g, k$  и  $\langle m1, m2, m3, m4 \rangle$  – значения сдвигов интенсивности кадров для выравнивания цвета фрагментов),  $A$  – алгоритм формирования обучающих пар и вектор входных признаков, который содержит интенсивность равномерной случайной выборки точек рассматриваемых кадров  $I1(\text{левый } t_i), I2(\text{правый } t_i), I3(\text{левый } t_{i+1}), I4(\text{правый } t_{i+1})$  на шаге  $n$ ,  

$$Ip(n) = \langle I1(x_i, y_i), I2(x_i, y_i), I3(x_i, y_i), I4(x_i, y_i), i = 1..256 \rangle$$
- нейронная сеть  $NN(Ip, P, W)$  на базе многослойного персептрона и алгоритма обучения градиентного типа, где  $W$  – множество параметров нейросети.

Основная задача такой модели в соответствии с особенностями конкретной стереосистемы камер проводить непрерывное обучение модели за счет поступающих кадров видеопоследовательности.

Алгоритм  $A$  – учителя включает в себя следующие шаги:

1. получение выборки  $Ip(n)$  на шаге  $n$ ,
2. вычисление начальных значений  $\langle m1, m2, m3, m4 \rangle$  как средних по фрагменту и первичное выравнивание фрагмента,
3. исполнение алгоритма предобработки с текущими параметрами,
4. проверка: если число формируемых сопряженных пар точек стереопары  $N \geq 200$ , то необходимо понижение порога  $k$  отсеечения по условию корреляции фрагментов  $(k - \Delta k)$ , повышение порога отсеечения фильтра края  $(g + \Delta g)$ , когда  $N \leq 100$ , то понижение порога отсеечения фильтра края  $g - \Delta g$  и повышение порога  $(k + \Delta k)$ , если  $200 \geq N \geq 100$ , то сохраняем текущие параметры для обучения и переходим к следующим кадрам,
5. если сделано больше максимального повторов или параметры  $g, k$  перешли границы допустимых интервалов, то проводим несколько циклов обучения со случайными вариациями

$\langle m_1, m_2, m_3, m_4 \rangle$ ;

6. возврат к шагу 3;

Процедура работы учителя может занимать значительное время в начале обучения и носит практически переборный характер. Первые циклы работы учителя служат для получения начальной выборки примеров и могут быть выполнены до начала движения. Общее время поиска «хороших» параметров  $P$  для каждой пары кадров в начале обучения составляет 5-10 циклов полной обработки кадров, что не позволяет реализовать режим реального времени. Однако по мере настройки сети начальные значения все больше соответствуют требованиям сцены и в большинстве случаев не требуют повторов обработки кадров. Общий сценарий работы СИЗ с обучением выглядит следующим образом:

1. Запуск алгоритмов СТЗ с пустой или сохраненной с предшествующих запусков базой примеров для обучения сети настройки параметров обработки кадров;
2. Получение кадров стереопар и вычисление вектора входа нейронной сети  $I_p(n)$
3. Определение текущего значения параметров алгоритма обработки  $P$  в нейросети (в пустой сети до начала обучения это практически случайное состояние выходов);
4. Запуск учителя для определения «правильных параметров  $P$ » на основе переборного механизма;
5. Сохранение подобранных параметров  $P$  и входного вектора  $I_p(n)$ ;
6. Обучение сети  $NN(I_p, P, W)$ , которое проводим только при накоплении значительного числа новых векторов в  $P$ ;
7. Возврат к п.2 или остановка обработки, если остановлен видеопоток.

Рассмотрим обработку тестовых примеров в сцене на рис.1. Процедура настройки параметров предобработки для пар рис. 1.б занимает 5 циклов при работе с начальной моделью параметров и только 2 подбора при работе с предварительно обученной системой. По итогам адаптации для этой системы кадров построено  $N=126$  сопряженных точек. Тестовый пример на рис.2 содержит 7 измерений движения камер СТЗ вдоль предполагаемой оси  $X$  системы (шаги по 100 мм, 1 налево, 5 направо и 1 такт стоит на месте). Общее перемещение составило 300 мм направо по оси  $X$  при пройденном пути 600 мм. Без включения модели адаптации финальный сдвиг определен  $\Delta X = 223mm$ , а с включением настройки

(после 200 тактов работы)  $\Delta X = 333\text{mm}$ , промах в первом случае составил 77 мм, а во втором 33 мм.

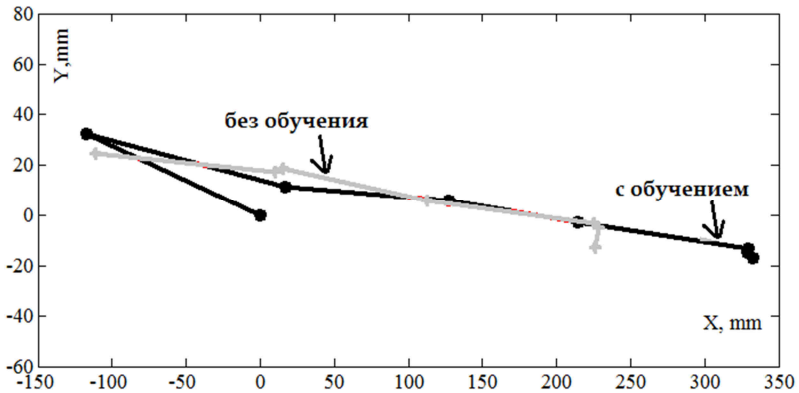


Рис. 2. Траектория движения СТЗ

Таким образом, очевидно, что некоторое усложнение модели обработки приводит к значительному повышению качества определения длины пройденного пути и координат положения беспилотной системы с навигацией на основе использования СТЗ.

## Заключение

Беспилотные мобильные системы требуют быстрой и интеллектуальной модели навигации. Такую навигацию могут обеспечить только многокомпонентные системы с использованием технического зрения.

Повышение быстродействия СТЗ возможно за счет использования непрерывно обучаемых нейронных сетей. Предложенная модель настройки СТЗ позволяет проводить выбор параметров предобработки кадров стереопары на основе обучения. Основными достоинствами такого подхода являются следующие моменты:

- решение будет приниматься на любом уровне обучения,
- с течением времени обучение будет улучшать качество настройки,
- время, затрачиваемое на выбор параметров на высоких уровнях обучения, сводится к прямому исполнению двухслойной нейронной сети,
- обучаемая настройка позволяет повысить качество результата.

Дальнейшая работа над моделью настройки СТЗ будет включать в себя

разработку механизма направленного поиска параметров выравнивания интенсивности фрагментов и реализацию ее бортового варианта.

## Список литературы

- [Девятериков и др, 2013] Девятериков Е.А., Михайлов Б.Б. : Управление движением мобильного робота с использованием данных визуального одометра // Робототехника и техническая кибернетика, №1,2013 стр.22-26.
- [Иванов, 2011] Иванов Ю.А. Технологии компьютерного зрения для наблюдения за объектами путевой инфраструктуры // Техника железных дорог, №4 (16), 2011, –С.57-61.
- [Жокарева и др., 2012] Жокарева 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
- [Корлякова и др, 2014] Корлякова, М.О., Пилипенко, А.Ю Нейросетевой поиск особых точек в стереопарах // Тезисы НТК “Техническое зрение в системах управления - 2014” , Москва , 18-20 марта 2014 - с. 125-127 [http://tvcs2014.technicalvision.ru/docs/сборник\\_тезисов\\_ТЗСУ\\_2014.pdf](http://tvcs2014.technicalvision.ru/docs/сборник_тезисов_ТЗСУ_2014.pdf)
- [Соколов и др, 2011] Соколов, С.М. Богуславский, А.А. Интеллектуальные алгоритмы обработки изображений для решения задачи распознавания в реальном времени бортовыми системами // Тезисы докладов научно-технической конференции-семинара. Техническое зрение в системах управления мобильными объектами. – М.: КДУ. – 2011. – С.86-88.
- [Сыряжкин и др, 2010] Сыряжкин В.И., Шидловский В.С. Корреляционно – экстремальные радионавигационные системы. – Томск: Изд-во Том. Ун-та, 2010. – 316 с.
- [Montemerlo et al, 2002] Montemerlo M., Thrun S., Koller D., Wegbreit B. FastSLAM: A factored solution to the simultaneous localization and mapping problem\ in Proc. AAAI National Conference on Artificial Intelligence, (Edmonton, Canada), AAAI, 2002. P. 593–598.
- [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