МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РФ

Федеральное государственное бюджетное образовательное учреждение высшего профессионального образования

«Тверской государственный университет»

Факультет прикладной математики и кибернетики

Кафедра информационных технологий

**ОТЧЕТ**

**по итогам научно-исследовательской работы**

**во 3-ом семестре**

Направление: *02.04.02 «Фундаментальная информатика и*

*информационные технологии»*

Магистерская программа: *«Информационные технологии в*

*управлении и принятии решений»*

Тема: «Сопоставление ключевых точек изображений с помощью нейронных сетей»

|  |  |  |
| --- | --- | --- |
|  | **Выполнил:**  студент 2 курса магистратуры  Иванов Артём Николаевич  **Научный руководитель:**  кандидат физ.-мат. наук,  Сорокин Сергей Владимирович  **Оценка:** | |
| Тверь – 2017 г. |  |

**ОГЛАВЛЕНИЕ**

[Введение 2](#_Toc502272604)

[Цели и задачи 2](#_Toc502272605)

[Актуальность 3](#_Toc502272606)

[Оценка состояния темы исследования 4](#_Toc502272607)

[1. Проделанная работа 5](#_Toc502272608)

[1.1. Задача сопоставления ключевых точек 5](#_Toc502272609)

[1.2. Разработка архитектуры нейронной сети 6](#_Toc502272610)

[1.3. Генерация обучающей выборки 9](#_Toc502272611)

[Заключение 11](#_Toc502272612)

[Список использованных источников 12](#_Toc502272613)

# **Введение**

В наше время очень много внимания уделяется совершенствованию технологий искусственного интеллекта. Большой популярностью пользуются нейронные сети, которые зачастую позволяют добиться лучших результатов[2], чем при использовании традиционных алгоритмов, как, например, в проблеме сопоставления ключевых точек изображений. Данная задача имеет широкий круг применения: создание панорам, создание стереопары и реконструкция трехмерной модели объекта по его двумерным проекциям в принципе, распознавание объектов и поиск по образцу из какой-то базы, слежение за движением объекта по нескольким снимкам, реконструкция аффинных преобразований изображений. Также следует отметить, что область знаний, рассматривающая такого рода задачи (компьютерное зрение) достаточно молода, соответственно еще нет определенного универсального метода, решающего все вышеперечисленные проблемы в полном объеме, т.е. для всех входных изображений. В тоже время существующие алгоритмы для решения данных задач постепенно совершенствуются, в том числе и с применением нейронных сетей.

## **Цели и задачи**

В данной научно-исследовательской работе поставлена цель:

* создать нейронную сеть для сопоставления ключевых точек изображений.

Для достижения данной цели необходимо выполнить ряд задач:

* ознакомиться с задачей сопоставления ключевых точек
* разработать архитектуру нейронной сети для сопоставления ключевых точек
* провести обучение и в случае необходимости доработку нейронной сети
* сравнить полученные результаты с известными методами
* оформить текст диссертации

В моей работе задача сопоставления ключевых точек рассматривается в контексте выделения структуры из движения (Structure from motion).

Это технология построения трехмерной структуры объекта по набору изображений (2D фотографий). Она является объектом исследования в таких областях как компьютерное зрение и зрительное восприятие (изучение зрительной системы головного мозга) [3].

Актуальность

Данная технология построения трехмерных структур используется во многих сферах деятельности: медицине, спелеологии, геологии. Существует множество инструментов для решения данной задачи. Наиболее популярными из них являются те, которые строят трехмерные структуры по видео или фото-потоку.

Специалистам требуется стоить трехмерные модели пещер и различных предметов местности. Существующие на сегодняшний день технологии имеют ряд недостатков. Так же показывают плохой результат работы в “специфических” условиях, к которым можно отнести резкий перепад света на последовательных кадрах.

Более того, при построении топологических карт местности данная технология по сравнению с аэросъемкой позволяет добиться большей точности при минимальном количестве затрат.

Также построение трехмерной структуры используется для оценки, мониторинга, представления и реставрации памятников культурного наследия[3].

Оценка состояния темы исследования

Существует множество алгоритмов для сопоставления особых точекна изображениях. Самыми известными и широко используемыми алгоритмами являются SURF (Speeded Up Robust Features) и SIFT (Scale-invariant feature transform) [1]. Некоторые из алгоритмов являются запатентованными, и показывают нерезультативность работы в сложных условиях. Недавние исследования по нейронным сетям предлагают использовать алгоритм LIFT, который превосходит результаты своих аналогов SIFT и SURF[2]. Но при этом всегда есть возможность улучшить скорость обучения и структуру сети для достижения более точных показателей работы.

# **Проделанная работа**

## **Задача сопоставления ключевых точек**

Так как моя задача решается в контексте Structure from motion, то рассмотрим поподробнее её реализацию:

* найти «особые» точки (точки интереса);
* определить соответствия между точками для пар изображений;
* отфильтровать ложные соответствия;
* найти трёхмерную структуру и положения камер.

Прежде всего дадим определение ключевой точке:

Ключевая точка — это такая точка изображения, которую можно отличить от любой другой точки изображения. На изображении можно выделить 2 типа особых точек: углы и ребра. Самым распространенным типом таких точек являются углы. Они, в отличие от ребер, не зависят от масштаба и расположения изображения[4]. Следовательно, их можно однозначно сопоставить для пар изображений. Углы можно определить, используя различные детекторы.

Детектор — это метод получения особых точек изображения. На вход детектору подается черно-белое изображение. На выходе получается матрица, значения элементов которой определяют степень правдоподобности нахождения угла в исследуемых пикселях изображения. После выполняется отсечение участков пикселей со степенью правдоподобности, меньшей некоторого порога. Оставшиеся после отсечения точки являются особыми.

Для нахождения соответствия между точками для пар изображений к каждой угловой точке должно прилагаться её описание (дескриптор). Если дескрипторы точек на разных изображениях близки, то можно считать, что один и тот же физический объект.

Дескриптор — это небольшая окрестность точки. В идеале, дескриптор должен быть независим от масштаба и ориентации изображения. Тогда поиск соответствий сводится к обходу всех дескрипторов одного изображения и поиску наиболее близкого дескриптора на другом изображении.

Для выполнения этого пункта существует множество алгоритмов: SURF, SIFT и т.д., но большинство из алгоритмов являются запатентованными, при этом показывают нерезультативность работы в сложных условиях.

## **Разработка архитектуры нейронной сети**

Искусственная нейронная сеть — математическая модель, а также её программное или аппаратное воплощение, построенная по принципу организации и функционирования биологических нейронных сетей — сетей нервных клеток живого организма. Это понятие возникло при изучении процессов, протекающих в мозге, и при попытке смоделировать эти процессы. Первой такой попыткой были нейронные сети У. Маккалока и У. Питтса.

Единицей данной сети является нейрон, модель которого схематически представлена на рис.1.



Рисунок 1 Модель нейрона

Совокупность данных нейронов образует сеть из параметров весов, которые регулируются в течение обучения нейронной сети: на вход ей даются данные из выборки, которые также содержат соответствующий идеальный отклик, который должна дать сеть на эти данные. После обработки входных данных, результат сети сравнивают с желаемым откликом. На этом этапе вычисляется функция потерь (loss function), которая высчитывает ошибку. После выполняется процесс обратного распространения ошибки - распространение сигналов ошибки от выходов сети к её входам, в направлении, обратном прямому распространению сигналов в обычном режиме работы.

Обращаясь к задаче Structure from motion, нам нужно получить нейронную сеть, которая будет сравнивать два дескриптора (описания окрестности точки изображения) и вычислять, описывает ли данный дескриптор одну и ту же точку, или же разные. Предполагаемая схема поиска соответствий представлена ниже на рис.2.

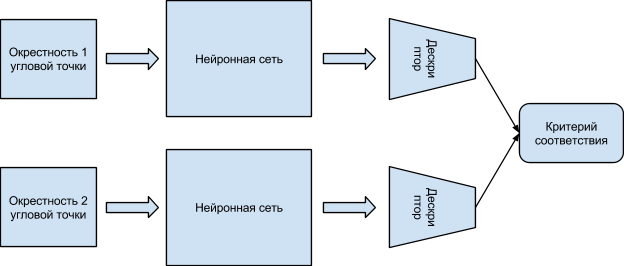


Рисунок 2 Схема нахождения соответствий

Как видим, для этой задаче лучше всего подходит модель сиамской нейронной сети – которая была впервые использована для нахождения соответствия между подписями группой ученых из AT&T Laboratories в США в 1994 году [5].

При этом, вместо стандартной реализации, использующей полносвязные слои перцептронов, будем использовать глубокие сверточные нейронные сети. Разработанная модель-прототип будет выглядеть следующим образом (рис.3):

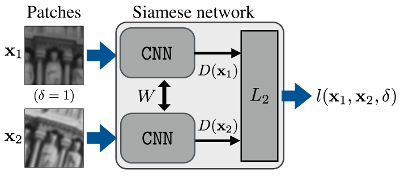


Рисунок 3 Модель-прототип

Как видно из рис.3, на вход подаются два окрестности точек изображения и параметр сигма, который указывает, принадлежат эти области одной и той же точке изображения (1) или разным (0). В основе сиамской нейронной сети лежит сверточная нейронная сеть, схема которой показана на рис.4.

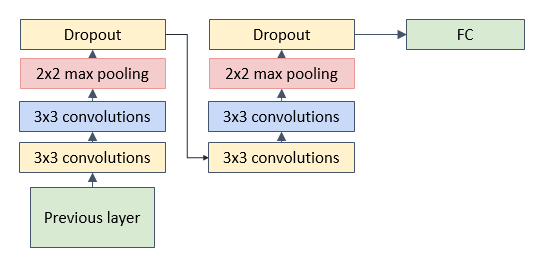


Рисунок 4 Схема CNN, входящей в сиамскую сеть

Это простое начальное представление модели, которую потом можно будет усовершенствовать. Она состоит из нескольких слоев сверток 3x3, слоев концентрации свойств (max pooling), слои с отсевом нейронов (Dropout) – для борьбы с переобучением. После каждого слоя свертки используется функция активации ReLU. Она также применяется после последнего полносвязного слоя (FC), который преобразует данные сети в одномерный вектор, который и будет являться дескриптором области точки.

## **Генерация обучающей выборки**

Задача обученной нейронной сети - проверять на равенство дескрипторы (фрагменты изображения с ключевыми точками), принимаемые на вход. Предлагается использовать следующий алгоритм для генерации обучающей выборки, который позволит с помощью поданного на вход изображения получить множество фрагментов с ключевыми точками с помощью аффинного преобразования изображения:

Алгоритм работы генератора:

1. Подаем на вход тестовую картинку A.
2. С помощью алгоритма из стандартной библиотеки opencv SIFT получаем набор координат особых (угловых) точек.
3. Поворачиваем картинку А на случайный угол, тем самым получая картинку B.
4. Берем случайную угловую точку из набора, полученного в пункте 2 находим соответствующую ей точку на измененной картинке B по формуле 2.
5. окрестности исследуемой точки на картинке A и картинке B, тем самым получая 2 фрагмента угловой точки размера N x N, где N может быть разным.

На основе данного алгоритма реализована небольшая программа для генерации выборки. Скриншот, показывающий её работу показан на рис.5.



Рисунок 5 Принцип работы алгоритма генерации обучающей выборки

Заключение

В этом семестре была проведена работа по изучению задачи сопоставления ключевых точек и области её применения – технологии Structure from motion. Также были изучены классический реализации нейронных сетей и их программная составляющая. На основе изученных данных была спроектирована модель-прототип нейронной сети для решения поставленной задачи работы. Вместе с этим написана программа по

В дальнейшем, согласно плану работ, будет произведено:

* программная реализация и доработка модели-прототипа нейронной сети и генерации обучающей выборки;
* обучение и доработка архитектуры нейронной сети;
* сравнение полученных результатов с известными методами;
* оформление текста диссертации.

Список использованных источников

1. А. Конушин, «Сопоставление изображений и локальные особенности», лекции по курсу «Введение в компьютерное зрение», 2012, стр. 45–79.
2. Сергей Белоусов, «LIFT: Learned Invariant Feature Transform»// Хабрахабр: интернет-сообщество индустрии высоких технологий [Электронный ресурс]. – Электрон. дан. – Режим доступа: https://habrahabr.ru/post/323688/
3. Structure from motion // Wikipedia, The Free Encyclopedia. December 27, 2017, [Электронный ресурс]. – Электрон. дан. – Режим доступа: <https://en.wikipedia.org/w/index.php?title=Structure_from_motion&oldid=816884383>
4. Обзор дескрипторов ключевых точек // НОУ «ИНТУИТ», [Электронный ресурс]. – Электрон. дан. – Режим доступа:

[http://www.intuit.ru/studies/courses/10621/1105/lecture/17983](http://www.intuit.ru/studies/courses/10621/1105/lecture/17983?page=2)

1. Jane Bromley, Isabelle Guyon, Yann LeCun, Eduard Sickinger, Roopak Shah, «Signature Verification using a "Siamese" Time Delay Neural Network»,1994