

# ОПРЕДЕЛЕНИЕ ПРОСТРАНСТВЕННОЙ ОРИЕНТАЦИИ ОБЪЕКТОВ ПО ИХ ИЗОБРАЖЕНИЯМ НА ОСНОВЕ НЕЙРОННЫХ СЕТЕЙ<sup>1</sup>

## DETERMINATION OF THE SPATIAL ORIENTATION OF OBJECTS FROM THEIR IMAGES BASED ON NEURAL NETWORKS

*V. Klyachin  
N. Peretyat'ko*

*Summary.* One of the tasks of 3D reconstruction of objects from their images is the problem of their spatial orientation relative to the coordinate system associated with the surveillance camera. A geometrically similar problem belongs to the class of inverse problems of projective geometry. However, the explicit solution of this problem is possible only in some special cases. Nevertheless, we set the task of determining the spatial orientation of buildings and structures from their photograph. The aim of the work is to solve the above problem by numerical methods, based on the presence in the photographs of buildings and structures of characteristic straight-line segments, which are the projection of lines running along the corners of the building. Geometrically, it can be justified that the presence of these segments is sufficient to solve the problem. Methods. Solving the problem is based on the use of a trained neural network. Novelty. Elements of the novelty of the presented solution are the use of image preprocessing for training by a neural network. Pre-processing consists in preliminary application of the Hough transform to select straight-line segments. Results. As shown by test calculations, the application of the proposed approach approximately doubles the accuracy of calculating the angle between the direction of the camera and the side plane of the building. Practical relevance. The practical application of the research results is to create a tool that allows you to automatically determine the relative position of objects on the photograph with their subsequent 3D reconstruction.

*Keywords:* projective geometry; spatial orientation; neural network.

**Клячин Владимир Александрович**

*Д.ф.-м.н., Волгоградский государственный  
университет  
klchnv@mail.ru*

**Перетьяко Никита Олегович**

*Аспирант, Волгоградский государственный  
университет  
vek2009@mail.ru*

*Аннотация.* Постановка задачи: одной из задач 3D реконструкции объектов по их изображениям является задача о пространственной ориентации их относительно системы координат, связанной с камерой наблюдения. Геометрически подобная задача относится к классу обратных задач проективной геометрии. Однако решение этой задачи в явном виде возможно только в некоторых частных случаях. Тем не менее, мы ставим задачу определения пространственной ориентации зданий и сооружений по их фотоснимку. Целью работы является решение вышеуказанной задачи численными методами, основанными на наличии на фотоснимках зданий и сооружений характерных прямолинейных отрезков, являющихся проекцией линий, идущих вдоль углов здания. Геометрически можно обосновать, что наличие этих отрезков достаточно для решения поставленной задачи. Используемые методы: решение задачи основано на использовании обученной нейронной сети. Новизна: элементом новизны представленного решения является использование предобработки изображений для обучения нейронной сети. Предобработка заключается в предварительном применении преобразования Хафа для выделения прямолинейных отрезков. Результат: как показали тестовые расчеты, применение предложенного подхода примерно в два раза повышает точность вычисления угла между направлением камеры и боковой плоскостью здания. Практическая значимость: практическое применение результатов исследования заключается в создании инструмента, позволяющего по фотоснимку в автоматическом режиме определить взаимное расположение на нем объектов с последующей их 3D реконструкцией.

*Ключевые слова:* проективная геометрия; пространственная ориентация; нейронная сеть.

### Актуальность

**М**оделирование по плоскому изображению, различного рода фото и видео материалам реальных объектов является наиболее сложной задачей 3D моделирования, особенно в задачах 3D реконструкции исторических архитектурных сооружений и целых комплексов, имеющих как историческую, так и культурную ценность. Задачи реконструкции культур-

ного и исторического наследия с использованием 3D технологий предполагают, в частности, решения ряда задач нацеленных на развитие методов и технологий 3D моделирования по плоскому изображению объекта при отсутствии какой-либо возможности доступа к реальному объекту. Это требует решения ряда геометрических задач, связанных с особенностями и условиями подобных 3D реконструкций. К таким задачам можно отнести: задачи определения характерных размеров реальных

<sup>1</sup> Работа выполнена при финансовой поддержке Российского фонда фундаментальных исследований и Администрации Волгоградской области (проект № 19-47-340015)



Рис. 1. Характерные элементы на изображении

объектов и их частей, задачи определения местоположения объектов в пространстве, задачи определения настроек камеры, с помощью которой было получено изображение, а также задачи моделирования параметрических прототипов объектов в виде геометрической 3D модели. Фундаментальная научно-техническая дисциплина, занимающаяся определением формы, размеров, положения и иных характеристик объектов по их фотоизображениям носит название фотограмметрия. Но надо сказать, что отличием классических задач фотограмметрии от задач рассматриваемых нами в данной статье являются неизвестные характеристики элементов внутреннего и внешнего ориентирования снимков (настроек камеры). В настоящей статье мы предпринимаем попытку применения нейронных сетей для автоматизации решения задачи нахождения этих характеристик. В основе предлагаемой методики лежит поиск характерных наборов точек, кривых и прямолинейных отрезков снимка, которые определяются существенно исходя из информации геометрического строения объ-

екта съемки. К сожалению в имеющейся литературе, например в курсе лекций [1], или в классической монографии [2] не удалось найти решение необходимых задач в требуемом виде. В частности, классические задачи фотограмметрии предполагают известными характеристики камеры (так называемые элементы внутреннего и внешнего ориентирования снимков), с помощью которой получено изображение. Тогда как в ситуации 3D реконструкции в задачах нашей работы это принципиально не известно и требует определения по характерным точкам и элементам на изображении, например таким как на рисунке 1. С целью решения задач определения тех или иных характеристик объектов по их снимкам имеется ряд публикаций, в которых решаются прикладные задачи в самых разнообразных отраслях науки и техники. Так в работах [3] — [5] изучается совокупность характерных точек на изображении, образующая размытость. А в работах [6],[7] авторы применяют сверточные нейронные сети для реконструкции поверхности лица человека по фотографии.



Рис. 2. Извлечение из изображения прямолинейных отрезков

### Постановка задачи

В настоящей статье предпринята попытка решения задачи определения пространственной ориентации объектов по их фотоснимкам. При этом сам объект 3D моделирования в настоящее время по тем или иным причинам не доступен для выполнения измерений. Соответствующая задача требует определения параметров формы объекта и аффинного преобразования проекции по некоторым геометрическим характеристикам его образов и прообразов: углам, размерам, площадям фигур, распределения оттенков цвета на изображениях. Решение этой задачи позволит определить реальные размеры и формы объектов по их изображениям или значения отношений характерных размеров этих объектов и их частей. Предполагается, что найденные решения могут быть использованы в задачах 3D реконструкции архитектурных комплексов довоенного Сталинграда. Описание соответствующей задачи было дано авторами в работе [8].

Поскольку в рассматриваемых задачах мы вычисляем параметры взаимного расположения камеры и объекта, то вместо определения параметров ориентации камеры в пространстве, мы будем определять параметры ориентации пространственных геометрических фигур, фиксируя положение камеры. Всюду в статье используется модель камеры, состоящая из трех уравнений центральной проекции

$$(x, y, z) \rightarrow (X, Y) = \left( \frac{x}{z} \delta, \frac{y}{z} \delta \right)$$

где  $(x, y, z)$  — точка в пространстве, величины  $X, Y$  моделируют пиксельные координаты на изображении, которое располагается в плоскости  $z = \delta, \delta > 0$ .

Предельным случаем центральной проекции является ортогональная проекция, заданная формулами

$$(x, y, z) \rightarrow (X, Y) = (x, y)$$

Основная решаемая в статье задача — вычисление ориентации здания по его фотоснимку относительно камеры съемки. Хорошо известно, что ориентация твердого тела в пространстве определяется тремя углами Эйлера. В настоящей статье мы ограничимся вычислением угла поворота вокруг вертикальной оси.

### Математическая основа для вычислений

В основе метода лежит геометрическая идея, состоящая в следующем. Рассмотрим в пространстве некоторый параллелепипед  $P$ . Рассмотрим три луча выходящих из некоторой его вершины и идущих вдоль его ребер. Проекции этих лучей образуют также три луча, выходящих из одной точки. В работе [10] были установлены формулы, позволяющие определить ориентацию параллелепипеда по величинам углов между проекциями рассматриваемых лучей. Определение на фотоснимке трех таких лучей в автоматическом режиме не представляет собой легкую задачу. Тем не менее, мы можем воспользоваться алгоритмами поиска всех прямолинейных отрезков методом преобразования Хафа и попытаться провести обучение нейронной сети, подавая на ее вход изображения этих отрезков (см. рисунок 2).

Слева на рисунке 2 показан результат определения прямолинейных отрезков методом преобразования Хафа на изображении справа. Мы можем видеть, что в этом наборе отрезков имеются те, которые являются проекциями прямых, которые в пространстве идут вдоль углов здания. Помимо этих отрезков имеются и те, которые были определены, но никак не связаны с нужными нам направлениями в пространстве. Наша нейронная сеть должна уметь отделить одни отрезки от других. Поэтому мы подготовим достаточно большое количество изображений, на которых нейронная сеть должна научиться выполнять такую фильтрацию.

Таблица 1. Структура CSV файла

Название файла изображения	Значение угла в радианах
./images/Building4/Building40.png	-0.5449942604108191
./images/Building4/Building41.png	-0.21274344745388446
./images/Building4/Building42.png	0.7165233159052476
./images/Building4/Building43.png	0.4351587551110086
./images/Building4/Building44.png	0.6944412618889607
...	...



Рис. 3. Примеры изображений 3D моделей строений

Подготовка набора данных и их предобработка

Подготовка данных для обучения нейронной сети выполнялась с использованием программы Blender. Имея в наборе ряд 3D моделей архитектурных сооружений и выполняя случайный их поворот вокруг вертикальной оси на угол в пределах от  $-45^\circ$  до  $+45^\circ$  делается рендеринг, а соответствующий файл изображения помещается в специальный каталог. При этом информация об имени файла изображения и угле поворота записывается в файл \*.csv. В таблице 1 показана структура такого файла.

На рисунке ниже представлены примеры получаемых изображений. В общей сложности было создано 12000 изображений.

После того, как файлы изображений подготовлены создается файл, в который помещается результат извлечения прямолинейных отрезков с этих изображений в виде матрицы  $135 \times 240$  значений пикселей в двух градациях цвета: 0 — черный цвет, 1 — белый цвет. Данные

упаковываются в виде списка таких матриц в файл формата \*.nprz с помощью библиотеки numpy.

Моделирование и обучение нейронной сети

При выполнении численных экспериментов использовалась библиотека keras на базе tensorflow для языка программирования Python и были выбраны два различных подхода. В первом случае нейронная сеть строилась как классификатор. В этом случае определялось  $m$  классов. Каждый класс с номером  $i$  определял диапазон углов от  $-\pi/4 + i \cdot \pi/2m$  до  $-\pi/4 + (i+1) \cdot \pi/2m$ ,  $i=0, 1, \dots, m-1$  разворота здания на изображении относительно прямой ортогональной плоскости камеры. Структура используемой нейронной сети показана на рисунке 4.

В качестве целевой функции (функции потерь) выбрана функция 'categorical\_crossentropy'. Значение  $m$  подбирается равным 30 (т.е. Точность предсказания  $3^\circ$ ).

Насколько предлагаемый подход лучше можно судить по результатам на рисунке 6. На этом рисунке

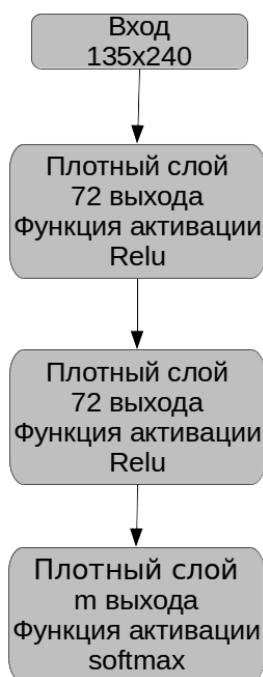


Рис. 4. Структура нейронной сети

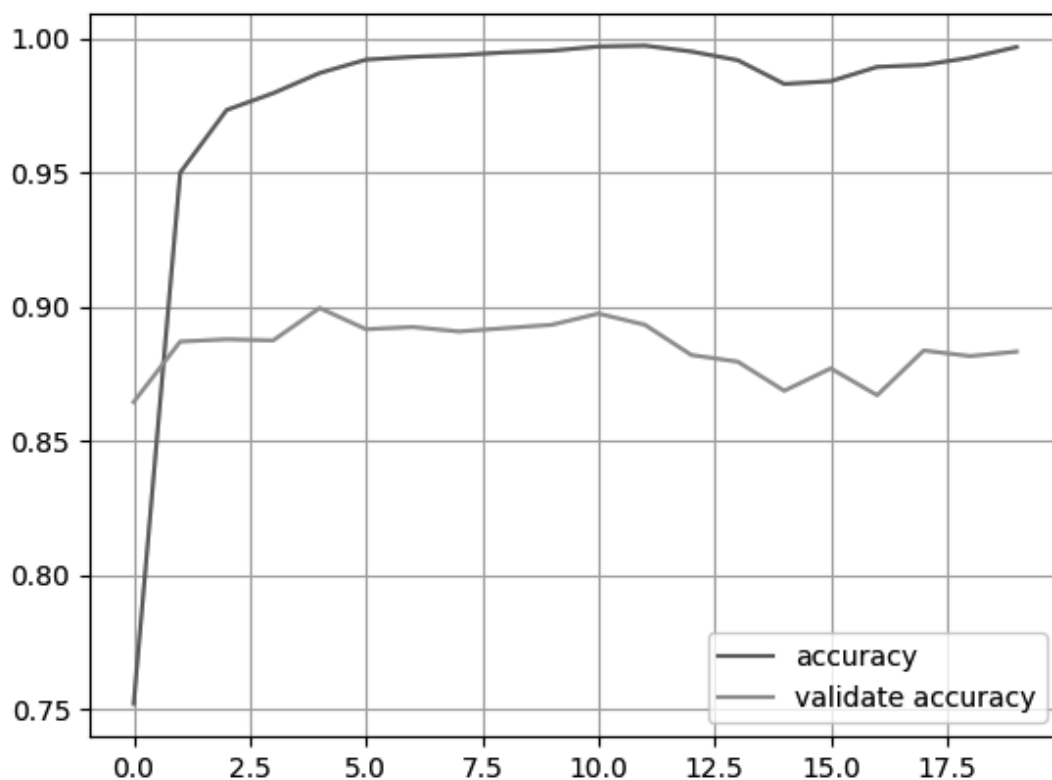


Рис. 5. Результаты обучения нейронной сети

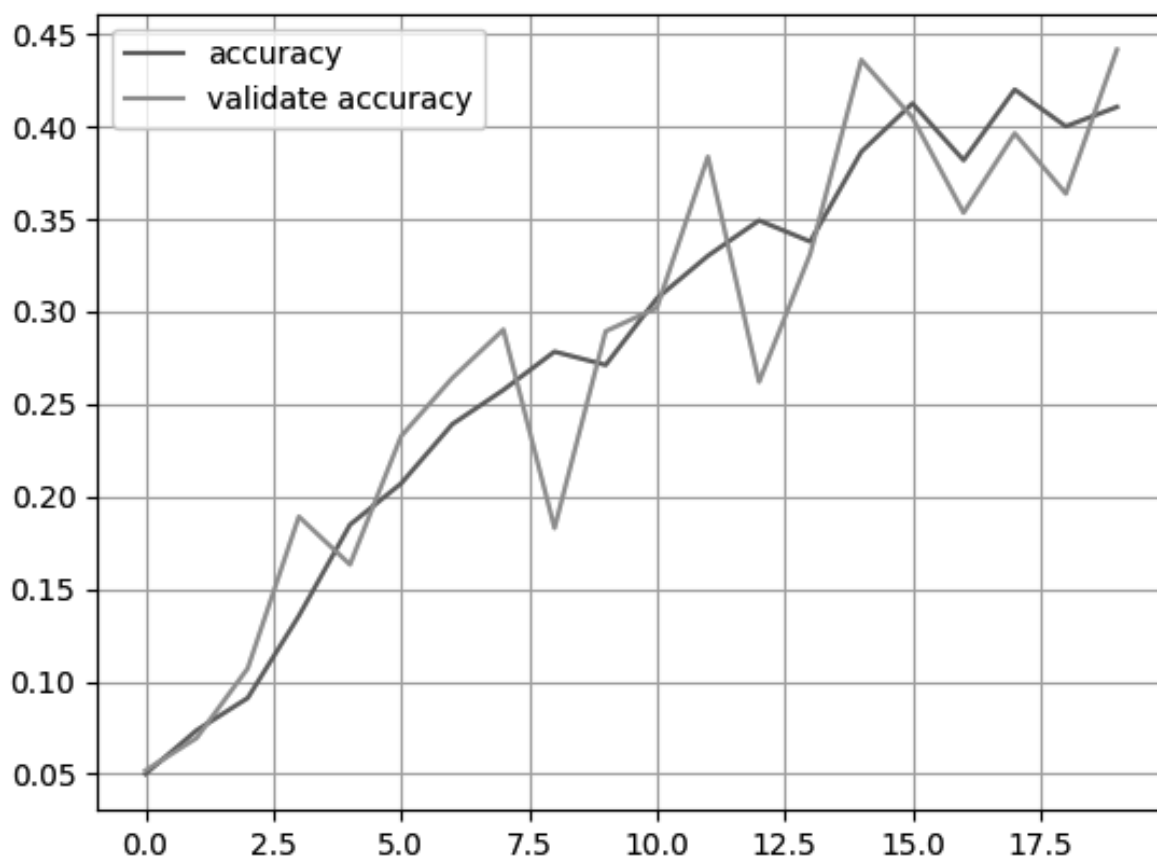


Рис. 6. Результаты обучения нейронной сети на исходных изображениях

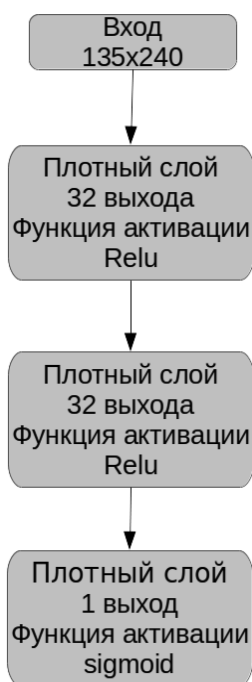


Рис. 7. Структура нейронной сети для регрессии угла

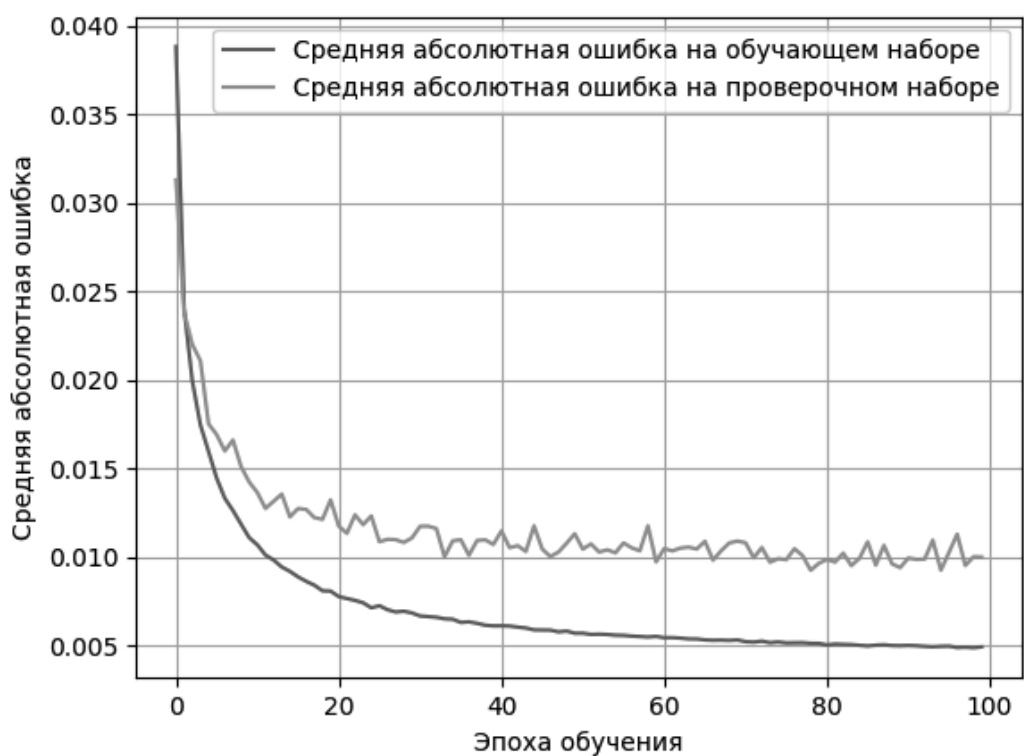


Рис. 8. Обучение регрессионной нейронной сети

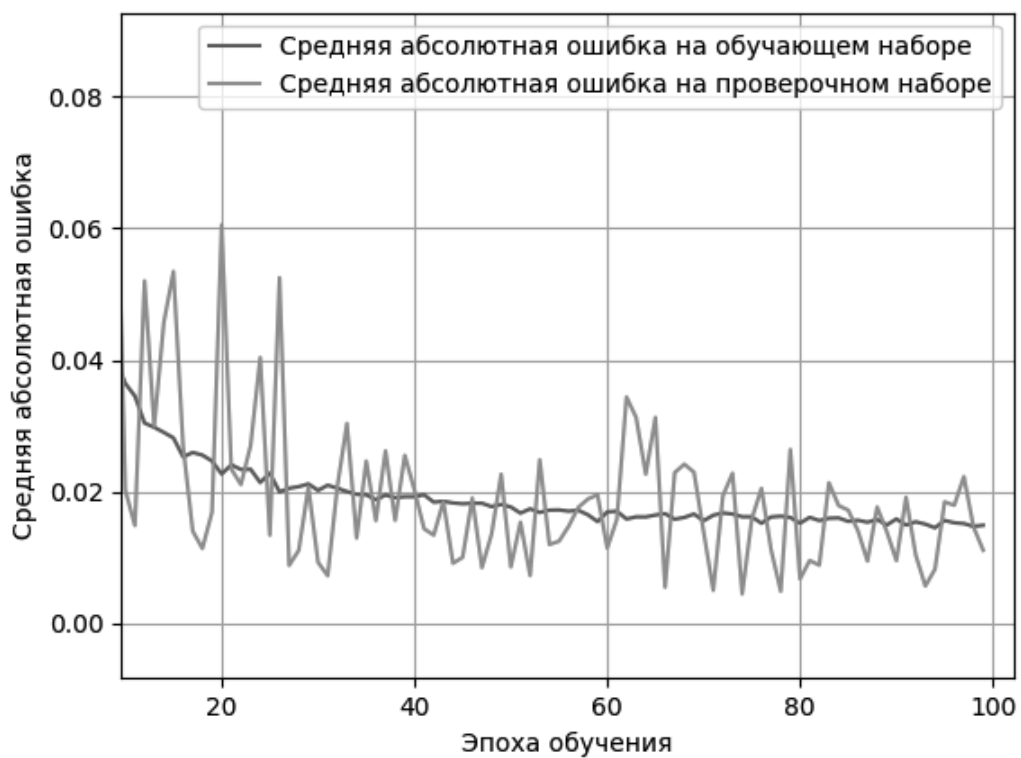


Рис. 9. Обучение регрессионной нейронной сети на исходных изображениях

показан процесс обучения нейронной сети с той же архитектурой, но когда вместо пред обработанного изображения (см. рисунок 2) в нейронную сеть подается непосредственно само изображение.

Рассмотрим второй подход, основанный на использовании функции потерь вида 'mse' (mean square error — среднеквадратичная ошибка). В этом случае нейронная сеть выполняет регрессию и предсказывает значение искомого угла. Поэтому в качестве меток мы используем не идентификатор класса, а непосредственно значение угла в радианах. Структура нейронной сети показана на рисунке 7.

Результаты тестирования показаны на рисунках 8 и 9.

Заметим, что как в первом случае, так и во втором точность вычисления угла на тестовом наборе данных

и на тренировочном наборе данных примерно в два раза лучше, чем при использование тех же нейронных сетей, но без предобработки изображений.

## Выводы

В статье предложен метод, основанный на применение предобработки изображений с помощью преобразования Хафа и построения нейронной сети для вычисления углов пространственной ориентации объектов зданий на фотоснимках. Показано, что точность вычисления повышается примерно в два раза по сравнению с ситуацией, когда предобработка не производится.

Работа выполнена при финансовой поддержке Российского фонда фундаментальных исследований и Администрации Волгоградской области (проект № 19-47-340015).

## ЛИТЕРАТУРА

1. Михайлов А. П. Курс лекций по фотограмметрии / Михайлов А. П., Чибуничев А. Г. — МИИГАиК. Ракурс (26 марта 2013).
2. А.Н. Лобанов Фотограмметрия / Н. Т. Куприна, З.Н. Чумаченко. — М.: «Недра», 1984. — 552 с.
3. Локтев Д.А., Быков Ю. А., Коваленко Н. Использование метода анализа размытия изображения для определения внешних дефектов железнодорожного пути // Наука и техника транспорта. 2016. № 1. С. 69–75.,
4. Локтев А.А., Локтев Д. А. Метод определения расстояния до объекта путем анализа размытия его изображения // Вестник МГСУ. 2015. № 6. С. 140–151.,
5. Локтев А.А., Бахтин В. Ф., Черников И. Ю., Локтев Д. А. Методика определения внешних дефектов сооружений путем анализа серии его изображений в системе мониторинга // Вестник МГСУ. 2015. № 3. С. 7–16.
6. Jackson A.S. et al. Large pose 3D face reconstruction from a single image via direct volumetric CNN regression //2017 IEEE International Conference on Computer Vision (ICCV).— IEEE, 2017. — С. 1031–1039.
7. Jackson A.S. et al. Large Pose 3D Face Reconstruction from a Single Image via Direct Volumetric CNN Regression //arXiv preprint arXiv:1703.07834.— 2017.
8. E. G. Grigorieva, V. A. Klyachin Mathematical model of 3D maps and design of information system for its control // Journal of Computational and Engineering Mathematics, 2016., № 4 p. 51–58.
9. OpenCV: Geometric Image Transformations [https://docs.opencv.org/3.4/da/d54/group\\_\\_imgproc\\_\\_transform.html](https://docs.opencv.org/3.4/da/d54/group__imgproc__transform.html)
10. Григорьева Е.Г., Клячин В. А. Алгоритм автоматического определения параметров камеры в пространстве на основе характерных элементов фотоснимка // Тенденции развития науки и образования. 2018. № 45–6. С. 10–20.