ДОПУСТИТИ ДО ЗАХИСТУ
Завідувач кафедри
В.М. Синєглазов
" 21 " 11 2022 р.

ДИПЛОМНА РОБОТА
(ПОЯСНЮВАЛЬНА ЗАПИСКА)

ВИПУСКНИКА ОСВІТНЬО-КВАЛІФІКАЦІЙНОГО РІВНЯ "МАГІСТР"

Тема: **Система визначення глибини зображення**

Виконавець: Сич О.С.

Керівник: к.т.н., доцент, **Василенко М. П.**

Консультант з екологічної безпеки: к.т.н., доцент **Явнюк А.А.**

Консультант з охорони праці: старший викладач **Козлітін О.О.**

Нормоконтролер: к.т.н., професор **Філяшкін М.К.**

Київ 2022

EDUCATION AND SCIENCE MINISTRY OF UKRAINE
NATIONAL AVIATION UNIVERSITY
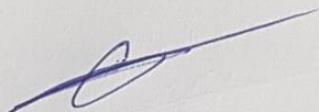## DEPARTMENT OF AVIATION COMPUTER INTEGRATED
COMPLEXES

ADMIT TO DEFENSE
Head of department
Viktor M. Sineglazov
" 21 " 11 _____ 2022

MASTER'S THESIS
(EXPLANATORY NOTE)

GRADUATE OF EDUCATION AND QUALIFICATION LEVEL
"MASTER"

THEME: Image Depth Detection System

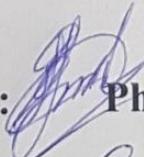Executor: Sych O.S.

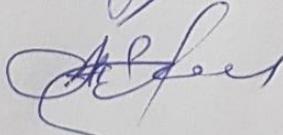Supervisor: Ph.D., Associate Professor Vasylenko M.P.

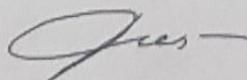Advisor on environmental protection: Ph.D., Associate Professor Iavniuk A.A.

Advisor on labor protection: Senior Lecturer Kozlitin O.O.

Norms inspector: Ph.D., Professor Filyashkin M.K.

Kyiv 2022

НАЦІОНАЛЬНИЙ АВІАЦІЙНИЙ УНІВЕРСИТЕТ

Факультет аеронавігації, електроніки та телекомунікацій

Кафедра авіаційних комп'ютерно - інтегрованих комплексів

Освітній ступінь магістр

Напрям 15 – Автоматизація та приладобудування

ЗАТВЕРДЖУЮ

Завідувач кафедри

Синєглазов В.М.

" 21 " 11 2022 р.

ЗАВДАННЯ
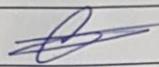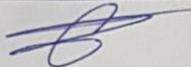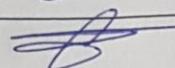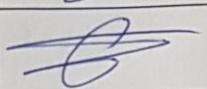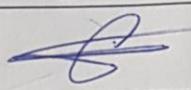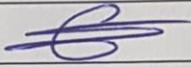### на виконання дипломної роботи студента
Сича Олексія Сергійовича

1. Тема роботи: «Система визначення глибини зображення».

2. Термін виконання роботи: з 19.09.2022р. до 15.11.2022р.

3. Вихідні дані до проекту (роботи): оптичний метод визначення дальності до об'єкта, карти глибини, алгоритми калібрування камер, фільтрація зображення, середовище Matlab.

4. Зміст пояснювальної записки (перелік питань, що підлягають розробці):

1. Актуальність системи визначення глибини зображення; 2. Огляд існуючих методів; 3. Огляд теоретичної інформації з приводу рішення задачі карти глибини; 4. Розробка системи визначення глибини зображення. 5. Вплив фільтрації зображення. 6. Вплив зовнішніх факторів на роботу системи.

5. Перелік обов'язкового графічного матеріалу: 1. Блок-схема калібрування камери; 2. Положення камер у просторі; 3. Структурна схема системи роботи з одним зображенням; 4. Карта розбіжностей у кольорі; 5. Таблиця перевірки точності щодо реального та виміряного діапазону за допомогою карти глибини в приміщенні та назовні; 6. Структурна схема пристрою виведення 3D-сцени.

6. Календарний план-графік

№ п/п	Завдання	Термін виконання	Відмітка про виконання
1	Отримання завдання	22.08.2022-23.08.2022	
2	Формування мети та основних завдань дослідження	24.08.2022-26.08.2022	
3	Аналіз існуючих методів	27.08.2022-31.08.2022	
4	Теоретичний розгляд рішення задачі	01.09.2022-08.09.2022	
5	Розробка структури системи оцінки глибини зображення	09.09.2022-15.09.2022	
6	Розробка програмного та апаратного забезпечення системи визначення глибини зображення	16.09.2022-30.09.2022	
7	Визначення та вплив методів фільтрації зображення на роботу системи	01.10.2022-12.10.2022	
8	Створення аналізу точності системи в різних умовах використання	13.10.2022-25.10.2022	
9	Оформлення пояснювальної записки	26.10.2022-30.10.2022	
10	Підготовка презентації та роздаткового матеріалу	01.11.2022-14.11.2022	

7. Консультанти зі спеціальних розділів

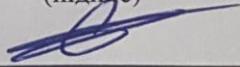| Розділ	Консультант (посада, П. І. Б.)	Дата, підпис	
		Завдання видав	Завдання прийняв
Охорона праці	Старший викладач, Козлітін О.О,		
Охорона навколишнього середовища	к.б.н., доцент, Явнюк А.А.		
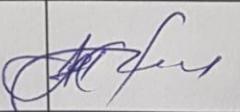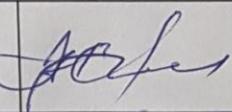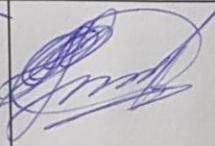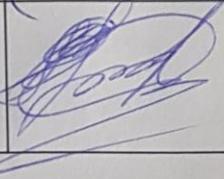
8. Дата видачі завдання   22.08.2022

Керівник: _____ Василенко М.П.
(підпис)

Завдання прийняв до виконання _____ Сич О.С.
(підпис)

NATIONAL AVIATION UNIVERSITY

Faculty of aeronavigation, electronics and telecommunications

Department of Aviation Computer Integrated Complexes

Educational level master

Field of study: 15 "Automation and Instrumentation"

<div align="right">

APPROVED BY

Head of department
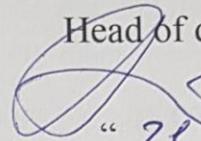
Victor M. Sineglazov

" 21 " 11 2022

</div>

Graduate Student's Diploma Thesis Assignment

Sych Oleksii Serhiyovych

1. **The thesis title:** «Image Depth Detection System».

2. **The thesis to be completed between:** from 19.09.2022 to 15.11.2022

3. **Output data for the thesis:** optical method of determining the distance to the object, depth maps, camera calibration algorithms, image filtering, Matlab environment.

4. **The content of the explanatory note (the list of problems to be considered):**

1. Relevance of the image depth detection system; 2. Review of existing methods; 3. Review of theoretical information regarding the solution of the depth map problem; 4. Development of a system for determining image depth. 5. Effect of image filtering. 6. Influence of external factors on system operation.

5. **List of compulsory graphic material:** 1. Block diagram of camera calibration; 2. Position of cameras in space; 3. Structural diagram of the system for working with one image; 4. Map of differences in color; 5. Accuracy check table for real and measured range using indoor and outdoor depth map; 6. Structural diagram of the 3D scene output device.

6. Planned schedule:

№	Task	Execution term	Execution mark
1	Receiving the task	22.08.2022-23.08.2022	
2	Formation of the goal and main tasks of the research	24.08.2022-26.08.2022	
3	Analysis of existing methods	27.08.2022-31.08.2022	
4	Theoretical consideration of problem solving	01.09.2022-08.09.2022	
5	Development of the structure of the image depth detection system	09.09.2022-15.09.2022	
6	Development of software and hardware of the image depth detection system	16.09.2022-30.09.2022	
7	Definition and impact of image filtering methods on system operation	01.10.2022-12.10.2022	
8	Creating an analysis of the accuracy of the system in different conditions of use	13.10.2022-25.10.2022	
9	Issuance of an explanatory note	26.10.2022-30.10.2022	
10	Preparation of presentation and handout	01.11.2022-14.11.2022	
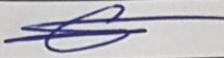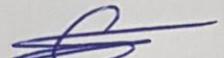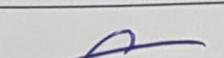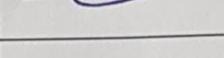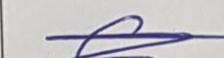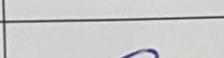
7. Special chapters' advisors

Chapter	Advisor (position, name)	Date, signature	
		Assignment issue date	Assignment accepted
Labor protection	Senior lecturer, Kozlitin O. O.		
Environmental protection	Ph.D, Associate Professor, Iavniuk A.A.		
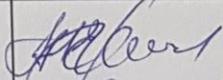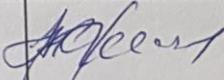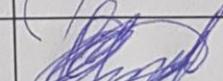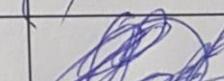
8. Date of task receiving: 22.08.2022

Diploma thesis supervisor _____ Mykola P. Vasylenko
(signature)
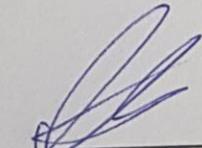
Issued task accepted _____ Oleksii S. Sych
(signature)
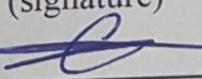
РЕФЕРАТ

Пояснювальна записка до дипломної роботи «Система визначення глибини зображення»: с., рис., табл., літературних джерела.

Об'єкт дослідження: процес оцінки глибини зображення.

Мета роботи: вдосконалення методів оцінки глибини зображення та розробки нової системи, що здатна з порівняно невеликими обчислювальними затратами здійснювати таку оцінку в різних середовищах використання.

Для досягнення цієї мети необхідно розв'язати наступні завдання:

- проаналізувати існуючі методи оцінки глибини зображення;
- проаналізувати існуючі апаратні засоби, придатні для реалізації розглянутих методів;
- на основі проведеного аналізу здійснити вибір методу який би дозволив здійснювати оцінку глибини зображення з мінімальними обчислювальними та фінансовими затратами;
- розробити програмне та апаратне забезпечення реалізації обраного методу;
- проаналізувати існуючи методи фільтрації зображення, для поліпшення результату вихідних даних;
- провести експериментальне дослідження роботи розробленої системи в різних умовах.

Предмет дослідження: розробка методу визначення глибини зображення для використання в різних умовах.

Методи дослідження: теоретична фізика, теорія оптики, теорія обробки зображення та машинного зору.

СТЕРЕО ЗІР; КАРТА РОЗРІЗНОСТІ КАРТА ГЛИБИНИ; КАЛІБРУВАННЯ; РЕКТИФІКАЦІЯ; 3D СЦЕНА.

ABSTRACT

Explanatory note to the thesis "Image Depth Detection System":   p., figures,   tables,   literary resources.

The object of research: image depth estimation process.

The purpose of the work: improving methods for estimating image depth and developing a new system that is able to perform such an assessment in various environments with relatively low computational costs.

To achieve this purpose, it must be solved the following tasks:

- to analyze the existing methods for estimating image depth;
- to analyze the existing hardware suitable for the implementation of the considered methods;
- on the basis of the analysis to make a choice of a method which would allow to carry out an estimation of depth of the image with the minimum computing and financial expenses;
- to develop software and hardware for the implementation of the selected method;
- analyze the existing image filtering methods to improve the result of the initial data;
- to conduct an experimental study of the developed system in various condition.

Subject of research: development of a method for detection image depth for use in various conditions.

Methods of research: theoretical physics, theory of optics, theory of image processing and machine vision.

STEREO VISION; DISPARITY MAP; DEPTH MAP; CALIBRATION; RECTIFICATION; 3D SCENE.

CONTENT

GLOSSARY

3D - Three-dimensional space

AI - Artificial intelligence

GPS - Global Positioning System

SDS-TWR - Symmetric Double Sided Two Way Ranging

RTT - Round Trip Time

ACK - Acknowledgement

PC - Portable Computer

CMOS - Complementary-symmetry/metal-oxide semiconductor

LiDAR - Light Detection and Ranging

USGS - United States Geological Survey

NOAA - National Oceanographic and Atmospheric Administration

NASA - National Aeronautics and Space Administration

IMU - Inertial Measurement Unit

GNSS - Global Navigation Satellite System

DNN - Deep Neural Network

RADAR - Radio Detection and Ranging

UA - Unmanned Aircraft

AC - Alternating Current

DC - Direct Current

WHO - World Health Organization

IARC - International Agency for Research on Cancer

TPP - Thermal Power Plant

MM - Measurement Method

TMPC - Technical Means of air Pollution Control

GNP - Gross National Product

CHP - Cogeneration or Combined Heat and Power

INTRODUCTION

In today's world, there is often a question about creating a model to solve a certain problem in such a way that it performs its intended task properly and does not have a large cost. This is what almost every developer of the project wants at the production stage.

Thus, the work consists in improving and acquiring better accuracy of the image depth detection system. For this, the model made for the bachelor's work was modified and improved, namely, the main design of the model was changed and the quality of the image was improved, thanks to various methods of image filtering.

Unlike the previous model, this project investigates the effect and quality of the 3D scene construction in the image, not the streaming video, under different weather conditions and at different observation points, which makes it possible to feel in more detail the impact of various phenomena on the model itself during operation and improve accuracy due to considering a single pair of images rather than a stream of large numbers at a specific frequency.

The design consists of two cameras, which were selected from the principle of price-quality, and a box to fix and protect the model itself, thus forming protection from the environment in various conditions of use. The design is connected to a computer that performs the software part, which consists in creating a stereo pair - artificial adjustment of cameras, image analysis at the initial stage and after filtering, which as a result gives an opportunity to see the difference in the accuracy of constructing a 3D image, which can be used for various goals, for example to find out the size or distance to the target object.

CHAPTER 1. RELEVANCE OF THE WORK

The space that surrounds people has a three-dimensional structure. When somebody want to ask a person to describe what person sees, it is not difficult to name the objects that surround.

Any object is primarily a set of information that is really available on the retina of the eye - it is no more, no less, a set of points, about a million. Each dot, or pixel, of an image simply contains information that gives an indication of the amount of light and color that comes from the surrounding space and is projected onto that spot on the retina. A phone, table or book does not exist on the retina. Guided by the information encoded in the image or the retina of the eye, and thanks to its own knowledge, the visual process builds perception. A phone or a book are the final answers as a result of the interpretive process that is an integral part of the visual system. Moreover, there is no term correspondence between sensory information, light and color and the response of objects. The visual system must provide the knowledge necessary to interpret and understand the external world.

The same algorithms can be built for a certain recognition system, which will in turn be considered as a computer vision system. So, it is necessary to consider what it is - computer vision.

Computer vision [1] is an area of artificial intelligence (AI) that enables computers and systems to extract meaningful information from digital images, videos, and other visual data, and to take actions or make recommendations based on that information. If artificial intelligence allows computers to think, then computer vision allows them to see, observe and understand.

Computer vision works just like human vision, except humans are one step ahead. Human vision has the advantage of being able to learn to distinguish objects,

determine their distance, know if objects are moving and if there is something wrong with the image.

Computer vision teaches machines to perform these functions, but it must do so in much less time with cameras, data, and algorithms than with retinas, optic nerves, and visual cortex. Because a system trained to inspect products or monitor a manufacturing asset can analyze thousands of products or processes per minute, spotting subtle defects or problems, it can quickly overwhelm human capabilities.

Computer vision is used in a variety of industries, from energy and utilities to manufacturing and automotive, and the market continues to grow.

In other words, computer vision, which is also a field of machine learning, studies the technologies and tools that allow computers to perceive and interpret information from the real world. AI gives machines cognitive ability, and cameras allow them to see. Together forming computer vision that allows to observe and understand.

Unlike computers, humans can naturally distinguish objects, judge the distance and understand the movement. On the other hand, the machine needs to be trained to perform these functions. Instead of eyes, it uses data and algorithms.

So the first part of the process is to give the computer "sight." To do this, need to connect the camera to the computer. The computer will then have to classify and interpret the collected data and determine its relationship and context. Therefore, his tasks are reduced to explaining what it sees in real time.

Computer vision is also a scientific discipline based on the theory of creating and using models and systems that are designed to extract information from images. The basis for research is not only static images, but also video data (sequence of images, images from cameras, three-dimensional data, etc.). This discipline is quite practical and technological, its theories and methods are actively used to create

computer vision systems for automated recognition of images and video data. Here are examples of such systems [2]:

- systems designed for process management (autopilots, industrial robots, automated control systems, machine vision);

- systems intended for video surveillance;

- systems designed for information processing and systematization;

- systems for modeling objects (in medicine: analysis of the results of microscopy, radiography, tomography, ultrasound examinations, etc.;

- in military affairs:

- analysis of topographical data, system of detection of enemy military forces;

- human-machine interaction modeling systems.

It is worth noting that this field is quite new and rapidly developing. It is generally accepted that computer vision began to actively develop as an applied field in the 70s of the 20th century, although there were certain developments in this area even earlier. The lack of real-world application examples has previously been explained by the lack of computers capable of processing large amounts of information. However, research in the field of computer vision has mostly been conducted within the framework of other scientific issues, so it is difficult to say that the specific problem of the discipline has been clearly formulated. But there are many point problems for which solutions have been formulated, although they are difficult to generalize to the scope of the discipline of computer vision. Quite a lot of solutions are currently at the level of theoretical development, various experiments are being conducted, etc. Along with this, more and more developed technologies have found application in real commercial products and are widely used in practice. In practical cases, it is possible to see that highly specialized

methods of computer programming are used for each task, which, however, are grouped on fairly generalized principles of computer vision. Within the framework of artificial intelligence tasks, the field of automatic decision-making and planning is important (for example, automatic robot movement, automation of warehouse equipment, etc.). For their effective implementation, it is necessary to receive information from the external environment about the object of decision-making by artificial intelligence. Computer vision helps in this by providing the necessary input information from video sensors, special sensors for the further operation of systems of interaction with the environment.

Also, computer vision is used for pattern recognition and learning applications, which belong to the field of artificial intelligence. Based on the above, it can be said that computer vision can be attributed to the subject area of artificial intelligence. In addition to computer science, this field is also closely related to physics. Many computer vision approaches are based on understanding physical phenomena, in particular, it concerns electromagnetic radiation (visible or infrared spectrum) reflected by objects. These radiations can be detected by special sensors and transmit the necessary information about the object. These processes are based on the principles of solid state physics and optics. To work with complex sensors and correctly read the image, it is necessary to understand the principles of quantum mechanics. Conversely, many physics problems can be solved using computer vision. As a result, computer vision can also be defined as a complement to physics. An equally important role is played by neurobiology in the study of computer vision, namely the understanding of the fundamentals of the functioning of biological vision. Currently, there is a lot of research on human and animal vision, the eyes of various mammals, the structure of the brain and the neurons responsible for recognizing objects, colors, etc. The principles of biological vision formed in this way became the basis for the development of computer vision systems, various systems were created that imitate the work of real biological systems at different levels of completeness. Analogues to one-dimensional signal processing methods

are widely used in computer vision and are extended to the analysis of two-dimensional and three-dimensional signals. However, there are methods of signal processing within the discipline of computer vision that have no analogues in standard one-dimensional signal processing methods (for example, they are distinguished by their nonlinear nature, multidimensionality, etc.). Therefore, signal processing can be considered under the domain of computer vision.

In general, many fields are related to computer vision (image processing, robotics, in the framework of robot perception of visual information, machine vision, etc.). However, all these areas have not yet been separated into a separate industry, although all the above areas are very closely related and have many common features. Currently, they are distinguished by the subject area, each of which has its own specifics. Let's give examples.

Image processing specializes in working with images in two-dimensional format. Its main task is image transformation, such as reducing contrast, adjusting edges, various transformations, noise removal, etc. At the same time, it is assumed that the performance of these actions does not depend on the content of the image itself.

Computer vision works in three-dimensional space. One of the main tasks is to restore the complete picture by one or more images. At the same time, the content of images plays an important role here.

Now that it has become clear what computer vision is, let's understand it in a little more detail. Indeed, browsing seems like an easy task. However, computer vision presents great difficulty in classification. In other words, it is not enough to recognize objects, there are other elements that must be taken into account in order to recognize them.

First of all, in the real world, virtual data is determined by the types of lighting. This first element can change the appearance or color of the object. In

addition, image noise can also make recognition difficult, as can object overlays. On the other hand, it is important to have an image from several angles so that the computer can recognize the object when the angle is changed, but one must start from the fact that the system should be as similar as possible to human vision. In addition, computer vision faces another difficulty: distinguishing objects of the same type and false similarity of objects at different distances. For example, there are many types of vehicles, but the algorithm must successfully recognize a truck and a motorcycle that are behind each other.

Taking into account all of the above, a system of image depth research was developed, with the help of which it is possible to find out the distances and sizes of this or that object. Currently, many analogs of this technology are used, but usually the price of such a system exceeds the expected result. The developed installation primarily performs the function of recognizing the position of the object in space. Let's highlight standard examples of recognition tasks:

- Image search by content: a specific unit with a defined content is selected from the array based on certain characteristics. The nature of the content can have different interpretations and selection parameters: similarity parameters, multi-criteria approach in text format, etc.

- Position estimation: determining the orientation of the object according to the position of the sensor, computer vision information reading camera. For example, determining the orientation of a product in a warehouse or on a conveyor line by a robot.

The problem of motion characterization is also popular. The speed of movement of certain points and their trajectory is determined on the sequence of images. The following tasks are solved:

- monitoring of camera movement;

- tracking the movement of the object.

In addition to object recognition and detection, computer vision can reconstruct images.

Image restoration can be interpreted as cleaning unnecessary noises (results of sensor defects, blurring, etc.). The simplest and most common solution to this problem is the application of certain filters of different frequencies. There are also more complex methods of solving this problem, based on the approach of determining "how it should be" and comparing it with "how it is". A more complex process of noise removal is based on the analysis of input data, the determination of certain structural elements, and then the use of more highly specialized filters based on the received information. Features of the practical implementation of computer vision systems strongly depend on the field of application. Systems can be separate functional solutions, or be part of a more complex structure and perform an auxiliary or intermediate function. The use of these systems can also vary depending on whether the tasks of the system are predetermined or can be supplemented or changed during its operation. As we can see, the functions of computer vision are quite diverse, but we can identify certain common points for all systems. Images are obtained using various sensors (light-sensitive cameras, radars, ultrasonic cameras, etc.). Images consist of pixels that usually express light intensity, less often other characteristics (absorption of electromagnetic waves, reflection of sound waves, depth, resonance, etc.). In order to use all the possibilities of computer vision methods, it is necessary to pre-process the image first. Example:

- contrast adjustment;

- selection of inaccuracies and noises and their elimination;

- checking the coordinate system;

- scaling.

Certain details can be highlighted in the image:

- lines;

- boundaries;

- corners;

- points;

- forms;

- movement;

- drops.

At a certain stage, those details of the image that will be used later in the work are highlighted. This process is called segmentation. Examples of segmentation:

- selection of the object;

- selection of a set of points;

- selection of contours, etc.

After preliminary processing and segmentation, we proceed to high-level processing. At this stage, the research object is already quite small, so it is possible to use more complex approaches for its analysis. Having made sure that there is an object of interest in the selected interval, and at this stage it is possible to determine its parameters: size, volume, color, position and others.

CHAPTER 2. ANALYSIS OF EXISTING ALGORITHMS FOR SOLVING THE PROBLEM

Today, there are many methods for determining the object's body and finding it in the image, which in turn makes it possible to learn more detailed information about it, namely its shape, dimensions, distance, and so on. In this section, the basic and advanced technologies used in the modern world will be considered.

2.1. Calculating the distance to an object from a photograph

Pattern recognition is a rather broad area, many problems have already been identified in this area. First of all, it is necessary to understand the basic concepts of this process. Main categories:

- A class is a set of objects that have common properties.

- Classification is the process of dividing objects into classes according to their characteristics.

- A classifier is a tool that receives a set of object properties as input, and outputs the class to which the object should be assigned.

- Verification is the process of matching a certain object with the object model or with the class description.

- A sign is a quantitative description of a property of the object or phenomenon being studied.

Therefore, it can be concluded that the task of object recognition is reduced to the following sequence of actions (Fig. 2.1).

ACIC DEPARTMENT		NAU 22 1094 000 EN			
Performed	O. S. Sych		N.	Page	Pages
Supervisor	M.P. Vasylenko	IMAGE DEPTH DETECTION SYSTEM		23	160
Consultant					
S. controller	M.K. Filyashkin		225 151		
Dep. head	V.M.Sineglazov				
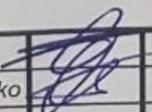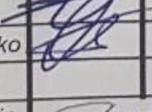
| Research, detection of input data (sensors, camera sensors, etc.) | → | Primary data processing (normalization, filtering, etc.) | → | Formation of feature vectors | → | Classification or prediction based on the received data |
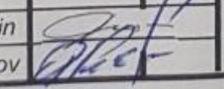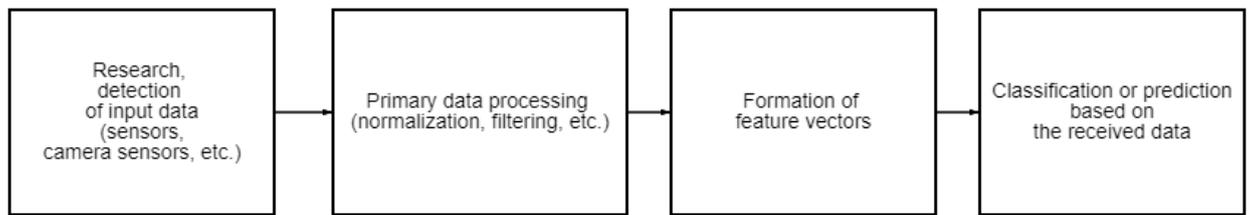
Fig. 2.1. Stages of object recognition

There are a number of methods and approaches for image analysis [3] and recognition, one of them is the method of contour analysis. The basis of this method is obtaining the external contour of the image of objects in the form of a closed curve or a collection of arcs. The points representing the contour of the image are recorded and represented using one of three methods for representing the boundaries of objects: curve approximation, contour tracing, or connecting points of difference.

This method is based on the assumption that the contour of the image contains the maximum necessary information about the shape of the object. It is important to note that the internal points of the object are not analyzed. On the one hand, this limits the functionality of the method and its flexibility, but on the other hand, it simplifies the calculation and algorithmic approach. The method of contour analysis allows to solve the main tasks of object recognition: moving, rotating, changing the scale of the object, etc. There is a study that notes human perception of images as a set of image formation of the contour of the object and its interior. From this point of view, this method is quite valid and justified. The main category of this approach is the concept of contour. Contour is a set of points (pixels) that separate the object from the background. A number of problems may arise when selecting a contour:

- insufficient change in brightness leads to a break in the contour;

- selection of extra contours due to the presence of noise in the image;

- wide contour lines due to the presence of noise, low-quality image, blurring, etc.

Initial point

Contour vector

1
1-*i*
1-*i*
-*i*
-1+*i*
-1+*i*
i
1+*i*

Contour pixel
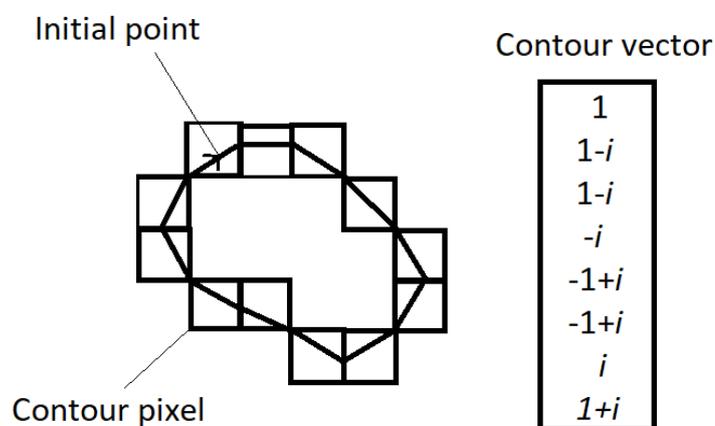
Fig. 2.2. Circuit coding scheme

Different approaches to contour coding (Fig. 2.2) can be used, the most famous among them: Freeman code, two-dimensional coding, polygonal coding. Contour coding during image recognition occurs using a sequence of complex numbers according to the following scheme:

- starting point is selected;

- the contour is described by elementary vectors that shift and are displayed using a complex number $a + ib$ (a – displacement of a point along the X axis, b – displacement of a point along the Y axis)

- the circuit closes when the end of the last vector coincides with the starting point.

There are several types of contour analysis:

- method of active contours (minimum energy curves are used to determine contours);

- the method of active contours without prior selection of borders (similar to the previous one except for the absence of the stage of preliminary selection of contours and image smoothing);

- Kenny boundary detector (the algorithm consists of smoothing, gradient search, suppression of non-maximums, double threshold filtering, tracing of uncertainty regions);

- tracing contours (the approach of "drawing" the border between the object and the background using a "bug" point);

- clustering;

- local processing (analysis of contour pixel characteristics);

- analysis using graphs (representation in the form of graphs and finding paths with the lowest cost corresponding to significant contours).

In general, the image recognition procedure using contour analysis can be described as follows:

- Pre-processing of the image (smoothing, filtering, noise separation, contrast enhancement, etc.)

- Image binarization and selection of object contours

- Initial filtering of contours by perimeter, area, fractality, shape factor, etc.;

- Reduction of contours to a single length, smoothing

- Going through all found contours, searching for templates that are as similar as possible to this contour.

The main advantage of this method is its relative simplicity. Filtering by simple characteristics is possible, and it is also possible to process images in progressive mode. Considering all this, we can conclude that this algorithm can be used in real time, which is an absolute advantage in the modern world. However,

the contour analysis method has a number of disadvantages. In general, we can distinguish two main groups of problems of this technique.

The first group is related to the difficulty of determining the contour in conditions of low image quality, presence of a significant amount of noise, lack of contrast between the background and the contour, etc. Thus, sometimes the algorithm cannot accurately determine the outline of the image, or it does it incorrectly. So it follows from this that the method of contour analysis is sensitive to image flaws, poor visibility is not allowed. Another group of problems is related to the principles of the method, namely, the algorithm does not take into account the possibility of intersecting objects, which is not always possible to achieve. Thus, the scope of its application can be significantly limited. In general, we can state that despite the existing shortcomings, the method of contour analysis is quite simple, fast, and therefore attractive for use in conditions of normal image quality and clear definition of object contours.

2.2. Creating an image depth map using AfterEffects software

First of all, nowadays many image processing programs have decided to add to their options such plugins or built-in functions for determining the depth or distance to objects. This is due to the advanced development of the fields of the game industry, cinematography, etc., AE is one such program.

The goal of the program is to process many templates of pre-recorded video, and to predict at what distance the object of interest is located by means of time shift simulations and changes. Also, the program is not entirely logical, due to the fact that the user himself can choose the point close to the camera and the far point, based on this, the question arises whether this will correspond to reality in the final result, because if you put in the foreground a conventional sky, which is located behind the object, for example a tree, and the tree will be in the background, the program will process this version of events, but the reality will be completely different.

This plugin is interesting in that it can adjust smoothing, filtering, image sharpness without much effort, and you can also change the frequency of the video frame stream itself.

The main disadvantage of such a program is that it does not make it possible to find out the exact distance or range to the body that is of interest, but only creates a certain model for the interests of the developer, it must also be said that this method is not fast in terms of processing, so when exposed maximum conditions and means of filtering, processing a five-second video can take more than an hour, this is even provided that the developer has a very good personal computer with excellent characteristics, which already creates a problem of price and whether it is worth doing it at all, if he the process will take a long time.

2.3. Conversion of an RGBD image into points in 3D space using the triangle method

This case will look at what the camera's internal matrix [4] is and how to use it to convert an RGBD (red, blue, green, depth) image into 3D space (Fig. 2.3, Fig. 2.4).
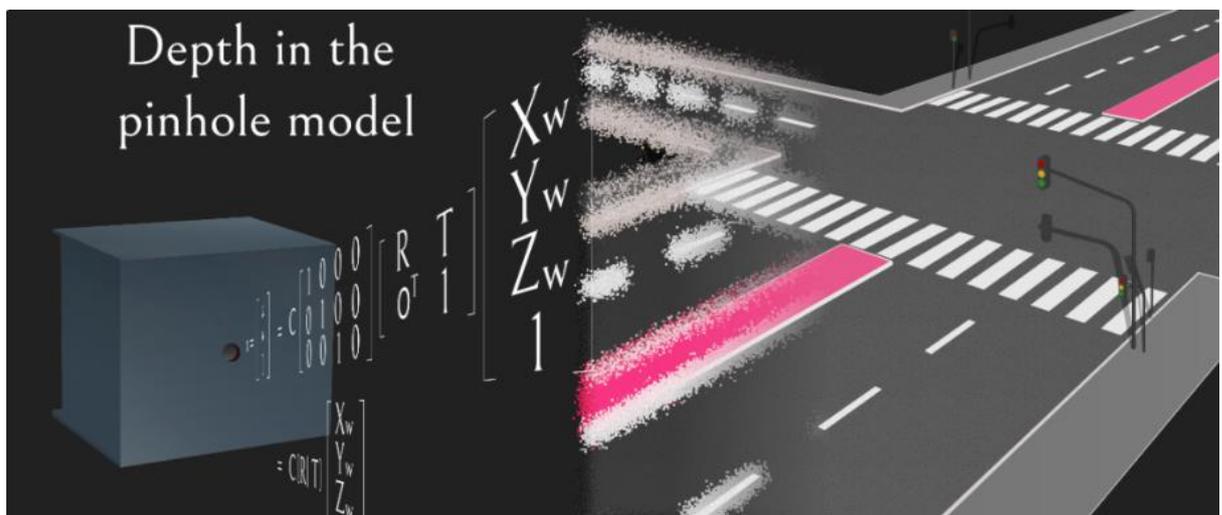
Fig. 2.3. A single-camera model of depth determination

Knowing the properties of the camera is very important if there is the need to understand what each pixel in the 3D environment corresponds to. One of the

key camera parameters is focal length, which helps convert pixel coordinates to actual length. It has been seen focal lengths such as "28mm", which is the actual distance between the lens and the film/sensor.

Using a simple geometric relationship ("similar triangles"), it is possible to easily obtain the x position from u and d for each pixel. The figure below only shows x and u, but they could be exactly the same for y and v. For the pinhole camera model, the focal length in the x and y directions is the same.
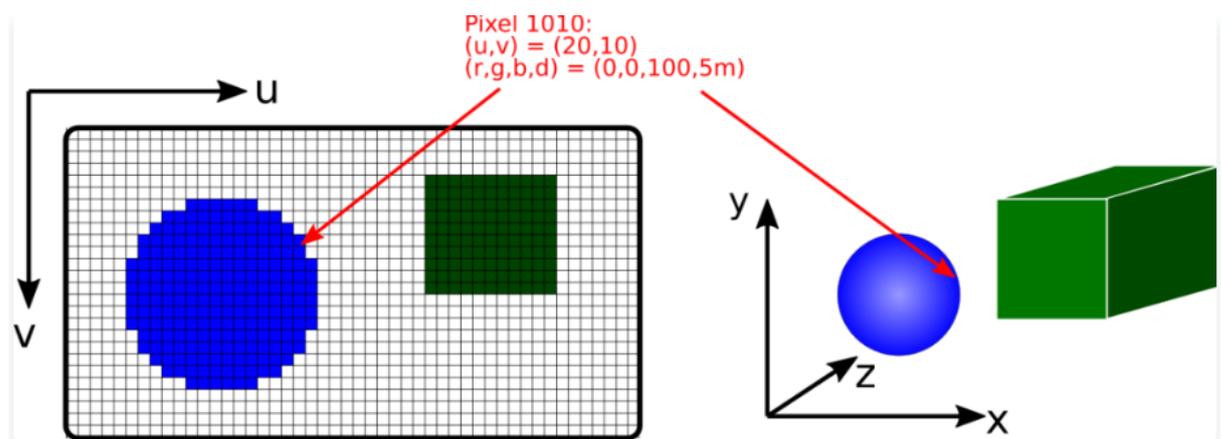
Fig. 2.4. (left) image plane in u,v coordinates. Each pixel has its own color and depth. (right) 3D view in Cartesian x, y, z coordinates

Knowing the properties of the camera is very important if want to understand what each pixel in the 3D environment corresponds to. One of the key camera parameters is focal length [5], which helps us convert pixel coordinates to actual length. Camera may have seen focal lengths such as "28mm", which is the actual distance between the lens and the film/sensor.

Using a simple geometric relationship ("similar triangles"), it is possible to easily obtain the x position from u and d for each pixel (Fig. 2.5). The figure below only shows x and u, but they could be exactly the same for y and v. For the pinhole camera model, the focal length in the x and y directions is the same.
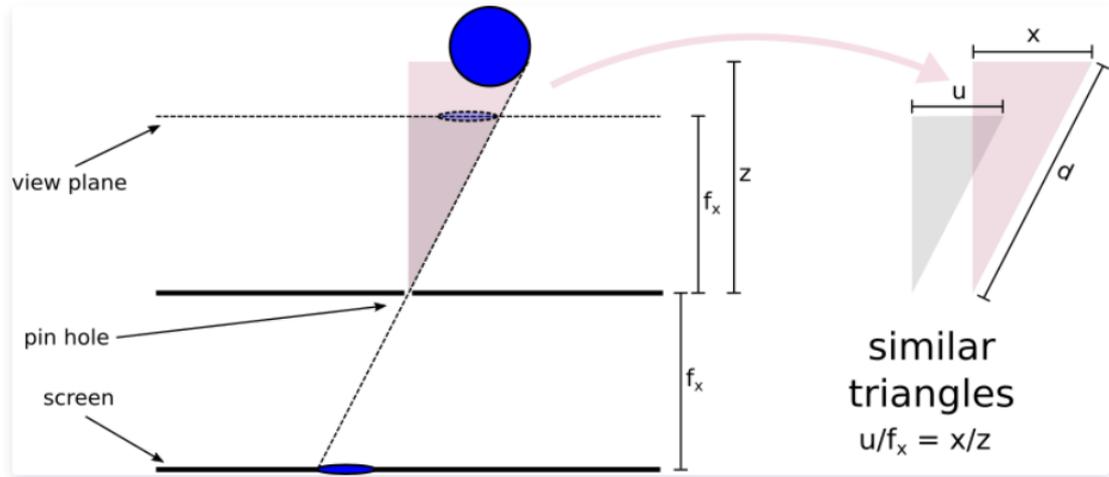
Fig. 2.5. Projection of the xz plane (top view)

On the left is the camera, in front of the lens (the same blue ball above) is the object displayed on the screen. The world coordinate system is aligned with the camera, so the z-axis extends in the direction the camera is looking. On the right, two triangles that partially overlap on the left are separated for clarity.

From a similar trigonometric approach, we immediately obtain:

$$x = \frac{u \cdot z}{f_x}. \tag{2.1}$$

Usually f_x and f_y are the same. But, for example, for non-rectangular image sensor pixels, lens distortion or image post-processing, they may differ.

In summary, it is possible to write a very short piece of Python code (Fig. 2.6) to convert the screen coordinate system to a Cartesian coordinate system using only geometric parameters.

```
1    def convert_from_uvd(self, u, v, d):
2        d *= self.pxToMetre
3        x_over_z = (self.cx - u) / self.focalx
4        y_over_z = (self.cy - v) / self.focaly
5        z = d / np.sqrt(1. + x_over_z**2 + y_over_z**2)
6        x = x_over_z * z
7        y = y_over_z * z
8        return x, y, z
```

Fig. 2.6. Code in the Python program

In the code (c_x, c_y) is the center of the camera sensor. Note the pxToMetre camera property constants, which can be used to accurately display 3D images at a specific scale.

Of course, there is a more general way to do all of this. Enter the matrix of internal parameters! This is a single matrix containing the camera properties discussed earlier (focal length and camera sensor center and skew). It is need to discuss how to use it to perform the above conversion. It then uses uppercase bold for matrices, lowercase bold for vectors, and regular font for scalars.

$$\boldsymbol{K} = \begin{bmatrix} f_x & S & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix}. \tag{2.2}$$

Next, introduce uniform coordinates. Homogeneous coordinates help to record various transformations (translations, rotations and skews) in the form of matrices of the same dimension.

In Fig. 2.5, can move the image plane any other distance, say from $f_x \rightarrow 2f_x$, and note the factor $h = 2$ that it moves to The shift introduces a simple scaling that can always be reversed by dividing u and v by h.

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \sim \begin{bmatrix} 2u \\ 2v \\ 2 \end{bmatrix}. \tag{2.3}$$

Now can define different operations on uniform coordinates, but these operations leave the value of the last dimension unchanged.

The rotation matrix R, the translation vector t and the eigenmatrix K form the projection matrix of the camera. Which is designed to convert world coordinates to screen coordinates:

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \frac{1}{z} \cdot \underbrace{K}_{3x3} \cdot \underbrace{|R|t|}_{3x4} \cdot \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}. \tag{2.4}$$

Note that $[R|t]$ denotes block notation, i.e. R and the column vector $t = transpose\{t_0, t_1, t_2\}$ are combined, in other words, we add it to the right side of R. If we want to do the transformation in another way, will encounter with problems, since the $3x4$ matrix is inverted. But it will turn out that the square matrix expansion of this matrix is reversible. To do this, one must add $1/z$ (parallax) to the left to satisfy the equation. A $4x4$ matrix is called an eigen/external matrix of full rank.

$$\begin{bmatrix} u \\ v \\ 1 \\ 1 \\ \frac{1}{z} \end{bmatrix} = \frac{1}{z} \cdot \begin{bmatrix} K & 0 \\ 0 & 1 \end{bmatrix}_{4x4} \cdot \begin{bmatrix} R & t \\ 0 & 1 \end{bmatrix}_{4x4} \cdot \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}. \tag{2.5}$$

In the simplest case, the origin of the camera is aligned with world coordinates, i.e. R and t are negligible, the skew S is 0, and the image sensor is centered. Now the reverse matrix of the camera:

$$
\begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = z \cdot \underbrace{\begin{bmatrix} 1/f_x & 0 & 0 & 0 \\ 0 & 1/f_y & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}}_{inverse\ with\ c_x, c_y, S=0} \cdot \begin{bmatrix} u \\ v \\ 1 \\ \dfrac{1}{z} \end{bmatrix}. \tag{2.6}
$$

Just looking at the first line, one can draw the same conclusion as at the beginning. The same is true for y and z in the second and third lines of equation. For more complex internal matrices, the inverse must be computed before performing this transformation. Since this is an upper triangular matrix, there are many easy ways to deal with this:

$$
\begin{bmatrix} \dfrac{1}{f_x} & -\dfrac{S}{f_x f_y} & \dfrac{S c_y - c_x f_y}{f_x f_y} \\ 0 & \dfrac{1}{f_y} & -\dfrac{c_y}{f_y} \\ 0 & 0 & 1 \end{bmatrix}. \tag{2.7}
$$

Now there are all the tools to convert a depth map or RGBD image into a 3D scene where each pixel represents a point (Fig. 2.7).
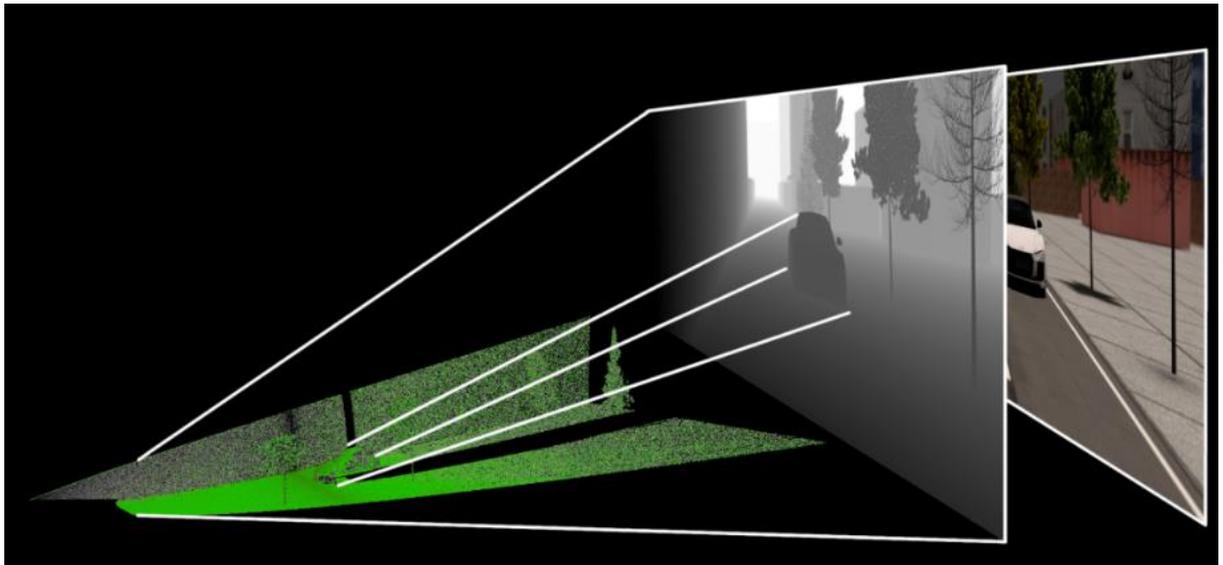
Fig. 2.7. Point cloud (green) calculated from the depth map (grayscale)

2.4 LIDAR

The LiDAR (Light Detection and Ranging) calculation system [6] has been created since the 1960s and works thanks to remote sensing technology, that is, it uses the results of calculations from laser pulses, which can then be used to build 3D models during processing, which in turn allows you to see a map of object and structure of the environment.

At the very beginning of the development of this technology, it was used on airplanes, and later, from the 1980s to be precise, after the appearance of less accurate GPS, the technology became more developed and improved to the point where it was able to calculate geospatial data.

The principle of operation of LiDAR is the same as that of radar or sonar, but if radar uses radio waves and sonar uses sound waves, then this system uses light radiation from a laser. To calculate, there is need to know the elapsed time of the laser to the reflection surface and back to the scanner, knowing the speed of light, it becomes very easy to find the real distance.

Since the speed of light is 299,792,458 meters per second, and the pulse generation frequency can reach more than 1,000,000 per second, then, when

calculating each of the points, it is possible to generate a cloud of points, which in turn will be used to construct a three-dimensional visualization.

Today, LiDAR technology is most often used to solve geodetic problems. Due to the fact that the system is able to collect a large amount of data to build a three-dimensional scene, it has been used to track artificial environments, namely, for the relief of a certain landscape (DEM) or the creation of digital terrain models (DTM).

Also, laser scanning is used to track natural phenomena, such as flooding, tracking coastal erosion, and the level of carbon accumulation in forests. The colossal impact of LiDAR has changed many applications and problems of automation of production processes. Cars began to use such systems for automatic control, of course with a reduced price and resolution of the scanner range, a shorter range, this is the system that, for example, the company Tesla used for its cars.

Cartographers also use the system to study the geographic and atmospheric position of objects on the map, for example, NASA, NOAA (National Oceanic and Atmospheric Administration), and USGS (US Geological Survey) have been using LiDAR for more than 15 years to study space and Earth maps.

Botanists, in turn, calculate changes in the structure and shape of forests on Earth.

Another interesting application is the analysis of gas compositions of the atmosphere. Since gases have different properties of absorbing light waves of different lengths, taking this into account, it is possible to calculate and determine which gas it is with the help of two laser beams of different lengths, which are then compared with each other by how much each of them was absorbed after reflecting from the surface. Over time, this system acquired the name DIAL - that is, LiDAR with differential absorption, it began to be used to measure changes in air pollution and find dangerous leaks in gas pipelines.

The most popular example of the use of LiDAR is the police car speed tracking device, although people are used to calling it radar, it is not.

Scientists and developers say that in the future, LiDAR technology will be advanced to extraordinary levels and will be in great demand, as they can be used to automate many areas of production: from 3D scanning to modeling and the development of smart cities.

With the advent of computer graphics, LiDAR began to be used to create and display virtual reality (LiDAR Augmented reality). The technology reproduces real content in the field of virtual reality with clearer graphics in the final result of an AR system. Because LiDAR performs high-quality "3D mapping," the point cloud data is overlaid on top of a regular data map, allowing for more realistic graphics and virtual reality components.

Research related to "Doppler winds" has arisen, which in the future will make it possible to investigate weather conditions that contribute to natural disasters, visualization of atmospheric data, problems related to aviation safety.

A laser scanning mapping system with an integrated Inertial Unit (IMU) and GNSS receiver is used in LiDAR Mapping. The points are combined with each other and form a map of the depth of the object or the field of view (Fig. 2.8).
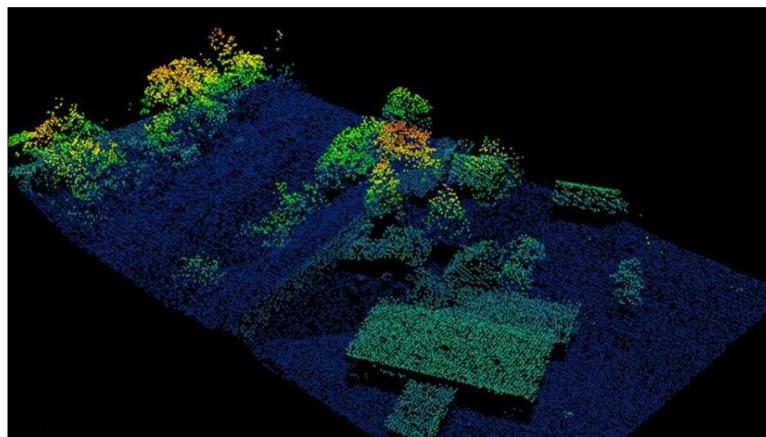
Fig. 2.8. 3D map image using LiDAR Mapping

These maps can be used to locate and determine the accuracy of the positioning of objects. In the future, point cloud data can be used to create a variety of high-precision depth maps of entire cities, buildings, roads for transport, and more, with an accuracy of up to a millimeter.

Advantages of using LiDAR technologies:

- Very accurate and fast collection of information;

- High penetrating ability;

- A person cannot contribute to the deterioration of the results, this is definitely a great advantage, since human error in measurements can cost too much;

- It is possible to implement other technologies to improve the result of information collection;

- There are no geometric distortions during movement;

- Good resulting image resolution;

- It does not depend on lighting, it works equally in daylight and at night.

Disadvantages of using LiDAR technologies:

- The cost is too high;

- Poor clarity of data collection in adverse weather conditions (rain, fog, snow, etc.);

- A very large amount of processing information that requires a good system to speed up the process;

- Gives poor results in conditions of water turbulence;

- It has a height limitation of operating characteristics, that is, when choosing a certain light wave, there is a risk that it will not be able to give information in a different environment at a different height.

2.5. Ultrasonic sensors

Just as LiDAR calculates the range to an object using a light wave, an ultrasonic rangefinder [7] uses a sound wave, just like bats or dolphins do. The principle of operation is built in such a way that pulses of a sound wave with a frequency of 40 kHz are emitted, an echo occurs, and during the propagation of the wave to the objects and back, it is possible to find out the distance at which the bodies are located.

When comparing ultrasonic and infrared rangefinders, ultrasonic comes first because it is not affected by problems related to lightning from sunlight or the color of the object's reflection, although there may be difficulties with reflections from fluffy or very thin objects, based on this, it becomes clear that it will be impossible to make high-precision calculations using an ultrasonic range finder.

The method of receiving and transmitting ultrasonic energy is used in many sensors and trackers. The frequency of which starts from 20 kHz, which creates an echo signal that is impossible for human hearing.

Several factors can contribute to the accuracy result:

- The distance at which the observed object is located;

- Ambient humidity and temperature;

- Class and manufacturing quality of elements of the module of a certain sensor;

- The position in which the object of observation is located relative to the sensor (related to the emission of sound waves).

As you already know, the principle of operation is based on the reflection of sound waves from the surface, but from the course of physics, it is known that the speed of sound propagation depends primarily on the density and temperature of the air itself. The sensor cannot adapt to the environment in which it is located,

therefore, the average value is usually taken for future calculations. Thus, sensors of this type can produce a measurement error of one to 5 centimeters.

The distance at which the object is located is also an element of error, since the farther the body is, the faster the sound waves will be attenuated.

A number of recommendations for improving the performance and accuracy of measurements have been identified:

- Multiple measurement with averaging (reduce to the average value from several measurements without taking into account surges during measurements);

- Addition of sensors to the system with the calculation of temperature and the introduction of a correction factor;

- Creation of a servo motor that allows you to change the orientation position to the right and to the left;

- It is also desirable to install the sensor opposite so that the object is within the directionality of the cone diagram.

The use of this technology is not entirely appropriate in many cases, as a number of problems described earlier often arise.

2.6. Infrared sensor

Another widely used device is an infrared sensor [8] that responds to changes in the background of thermal radiation in the area of use. Each of the surrounding bodies has its own thermal background, regardless of whether it is a person or not. The main purpose of the sensor is to highlight the object of observation against the general background of the heat map and react to body movement.

Usually, after a body of a certain size and temperature begins to move in the serviced area, a signal is sent to the control panel, which, in turn, then turns on certain automation, for example, a light or a security system. Since the sensor cannot cover a large area for observation, two or more devices are usually installed.

Modern gaming technology Kinect (Fig. 2.9) uses just such principles.

Fig. 2.9. Using Kinect technology on TV

Project Natal or Kinect [9] in modern terms is a non-contact touch-type game controller, the first use was in the Xbox 360 console, then Microsoft began to develop this field and extended the life of Kinect for further Xbox models and personalized programs for OC Windows. Thanks to an additional device installed specifically for interaction with a person, body movements are monitored, which are transmitted to the processing system.

A monochrome CMOS matrix is built into Kinect, which helps the depth sensor to get a good three-dimensional image regardless of room lighting. Independent calibration, which is adjusted automatically, regardless of the movement of furniture and things in the room, is also a good plus.

The main disadvantage is that this system was from the very beginning and until now, developed only with exclusive software, which increases the cost in case of modification of the design itself by the developer.

2.7. Analysis of the method of calculating distances to objects using one photograph

Each ordinary camera has its own characteristics, but one of the most important of them for solving the problem of calculating distances is the focal length. If the camera has a lens and is not a pinhole camera (usually the front part of such devices is called a pinhole camera), then the simplest option for creating a lens is to use one biconvex lens. The method of solving the problem using one lens [10] will be considered, then more advanced cases.

As already known from the usual course of physics, namely geometric optics, it will be considered the scheme of the course of thin rays using the example of a thin lens (Fig. 2.10).

Where f is the focal length of the lens itself, d is the distance of interest, namely from the object to the lens, D is the distance from the lens to the projection matrix or image film.
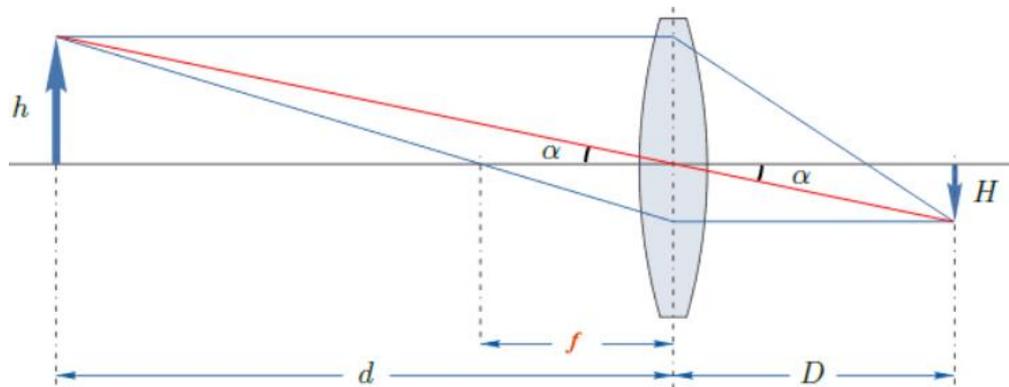
Fig. 2.10. The scheme of the path of the rays on the example of a thin lens

The same formula of a thin lens looks like this:

$$\frac{1}{d} + \frac{1}{D} = \frac{1}{f}. \tag{2.8}$$

Let's consider the optical scheme in more detail: H is the size of the projecting or reduced image on the matrix or film, h is the size of the observation object on the image, as already understood $H = D \cdot \tan \alpha$, and $h = d \cdot \tan \alpha$ (follows from properties of right triangles). After using the substitution, the modified formula becomes:

$$1 + \frac{h}{H} = \frac{d}{f}. \tag{2.9}$$

Since the formula got rid of, in most cases, the unknown variable D, it becomes easy to derive the formula for calculating the distance to the body or object of observation:

$$d = \frac{f \cdot (H + h)}{H}. \tag{2.10}$$

Since the formulation of the basic calculation formula is completed, it can be moved on to the example. Let the body, which will be photographed by the camera, be a building of a certain size (Fig. 2.11), which can be seen in its entirety. From what has already been described, it is possible to understand that a number of variables that represent interest, in connection with the formula described above, can be known. Three variables - f, H and h - will be needed for accurate calculation. The size of the house h is unknown in advance, but knowing a certain list of information, it is possible to calculate it. For example, if the average height from the floor to the ceiling is 2.64 meters, and the thickness of the overlap between

the floors is about 0.22 m. Rounding the value and taking into account the height of the floor covering, the resulting value is somewhere around 2.9 m, which is not a completely accurate characteristic, if it is possible to get a plan of the building, then the obtained value will be many times more accurate. The photo shows that the building has 23 floors, so the total height is approximately 66.7 m.

The method of finding size H, that is, the size of the display matrix on the camera, must be calculated, knowing the technical characteristics of the device, for example, the Nikon D90 camera is considered, the resolution of which is 4288×2848 pixels, and the screen matrix itself has a size of 2.36×1.58 centimeters. In order to find the size of the projection on the matrix, the image must not be cropped or rotated, after which you will need third-party software, such as Adobe Photoshop.

Fig. 2.11. Photo of the building

To calculate, there is need to make several mathematical transformations, the first of them is the transformation of the matrix into centimeter units of

measurement $4288/2.36 = 1817\ pixels/cm$, which is the correct resolution. The imprint size row is already filled with a value of 2.36×1.57 centimeters. Although, according to the indicated figure in the specification, it should be 1.58 cm, it is an insignificant error for further calculations. Using the built-in Photoshop function "Ruler", the height of the side of the house that is more visible to the observer is measured (Fig. 2.12).
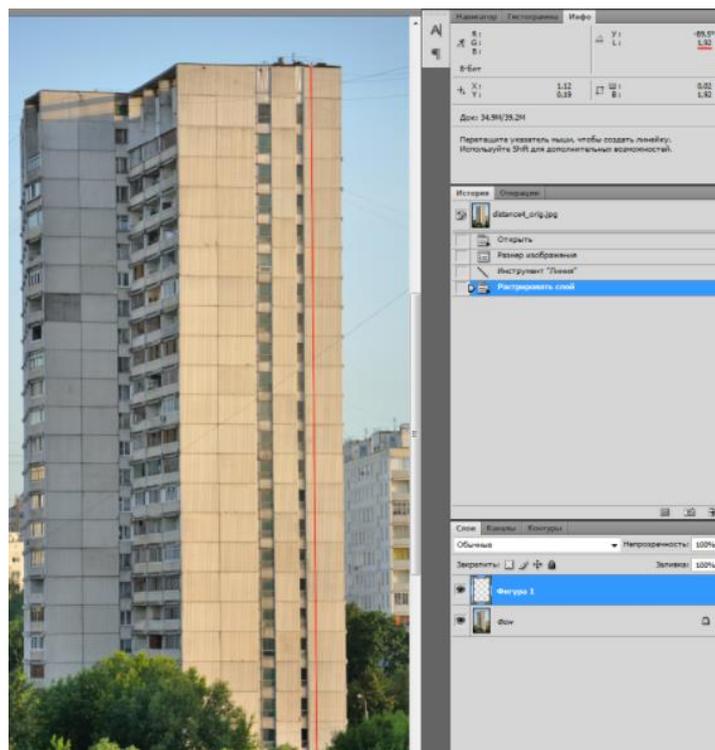
Fig. 2.12. The most visible edge of the height of the building

The resulting value is 1.92 centimeters or 0.0192 meters of height projection on the matrix. Using the same Photoshop, it is possible to find out the focal length, which in this frame has a value of 105 millimeters or 0.105 meters. Since all unknown variables are accounted for, the result can be calculated using formula 2.10:

$$d = \frac{\left(0{,}105 \cdot (0{,}0192\ +\ 66{,}7)\right)}{0{,}0192} = 364{,}9\ \text{m.}$$

This method is good considering that only one camera and certain characteristics are needed to determine such a task, but it is suboptimal in cases of calculating the distance to objects that are closer than 10 meters and cannot be fully focused, due to small details that deteriorate the quality of the display, also the most important drawback is that without previously obtained data about the size of the object, the use becomes impractical, since it will only be guesses about the real dimensions.

2.8. NanoLoc technology

Nanolock technology [11] calculates and determines the signal between the transmitting and receiving device, which in turn allows to automatically determine the distance with an accuracy of one meter. This technology was developed on the basis of the product of the production of systems using the Time of Flight method, and later acquired the name Symmetric Double Sided Two Way Ranging (SDS TWR), which also used the Round Trip Time (RTT) method.

RTT works on the principle of measuring the time of sending and receiving ACK packets, that is, object A sends a signal to point B, fixing the sending time, point B receives the signal and, after fixing the arrival time, sends a return signal of ACK packets. At the end, the object, having accepted the final option and recording the arrival time from object B, starts processing the information. What is interesting about this method is that the processing of signals in both objects must take place in the same time, and the fixing of the time of departure and arrival is done by hardware. Thanks to this, it is possible to predict the predicted propagation time of the signal according to the formula:

$$t_p = \frac{T_{RTT} - T_{reply}}{2}, \hspace{3cm} (2.11)$$

where T_{RTT} is the time of sending from object A to B, and T_{reply}, on the contrary, from B to A before sending ACK packets.

Since the speed of signal propagation is considered known, it is not difficult to calculate further data on the distance between objects. The quality and accuracy of the NanoLoc technology directly depends on the tracer modules, which must have a stable crystal frequency. Accuracy errors are described by the dependence part per million (ppm), which have a ratio of $1\ ppm = 0.0001\%$ error. To explain the calculation of errors, an example is considered in which the frequency of the crystal is $4\ MHz$ in the nominal value, while the stability frequency is $1\ ppm$. Thus, the deviation error from the nominal frequency will be $4000000 \cdot 10^{-6} = 4\ Hz$.

After testing the NanoLoc technology [12] by the researchers, it was found that the system has very good but sometimes overestimated measurement accuracy results, which directly depends on external conditions. Problems related to deviation of accuracy due to external conditions can be solved by early calibration of the system or by adding the influence of graduated curves. In most cases, the error is caused by nearby objects or any obstacles from which the signal can also be repelled, errors also occur due to the incorrect arrangement of antennas that receive and send signals.

2.9. Radar method

The field of radio electronics, which deals with determining the position, observing an object, calculating the coordinates of a position in space, physical dimensions or movement of a certain body using various methods and means of radio engineering, is called radar. The problems and tasks that were listed above are solved thanks to radars or radar stations.

The use of radar [13] has spread in many areas of use: detection of atmospheric formations, unmanned (UA) and manned aircraft, artificial and natural

space bodies, ships of various types, underwater, surface and underground objects, etc.

For example, in aircraft, the main task of the radar is radar sounding, determining and obtaining information about coordinates in space (angular coordinates and range to the target).

In the scientific world, the process of determining angular coordinates using radar is called radio direction finding, and ranges are called radio ranging. Before you can find out the speed or coordinates of the target, you need to identify and define the object. Recognition is the definition of, for example, significant distinctions based on certain characteristics, such as relation to a certain state. Identification is the process of processing the received radar data, which is carried out during target recognition.

After receiving data from radar devices, which is called radar information, the data is sent to control devices, command posts or calculation and solving devices.

The principle of operation of receiving data from radar devices is built on the basis of electromagnetic waves, which act as carriers for radar signals.

Electromagnetic waves at a constant speed propagate in a straight line in a homogeneous medium, according to the formula:

$$v = \frac{1}{\sqrt{\varepsilon_a \cdot \mu_a}}, \qquad (2.12)$$

where μ_a is the absolute magnetic and ε_a is the absolute dielectric permeability of the radio wave medium.

The basis of radar measurements depends on the directivity and propagation modulus of the stability of the propagation velocity vector in a homogeneous electromagnetic wave environment.

To find out the distance between objects, there is need to know the time of radio wave travel between the radar system and the target t_d. Thus, the range to the target is calculated according to the formula:

$$D = c \cdot t_d. \tag{2.13}$$

Due to the fact that there may be cases of inhomogeneity of the free space, which occurs due to a violation of the radio wave propagation speed vector, since there is no coincidence of the parameters ε_0, μ_0 and ε_a, μ_a. This can contribute to the modification of radio radiation, that is, there can be cases when some part of the energy is refracted, absorbed or re-reflected from the object. If it is considered the first of the options, that is, when the signal is refracted and becomes a source of secondary radiation, then the time of the reflected signal relative to the emitted signal with a delay is calculated by the formula:

$$t_d = \frac{2 \cdot D}{c} . \tag{2.14}$$

Taking this into account, the slant range to the target is calculated:

$$D = \frac{c \cdot t_d}{2}. \tag{2.15}$$

Another calculation option with greater accuracy is a radar with an additional repeater, which processes the sounding signal of the radar system and transmitter launch, and then receives the corresponding signal, the range to the target is calculated according to the formula:

$$D = \frac{c \cdot (t_d - t_{ans})}{2},$$ (2.16)

where t_{ans} is the delay time of a previously known signal in the responder circuit, and t_d is the delay time of the signal in response to the probe.

The time t_d is measured using special inertia less clocks of the electronic type, since the delay time of the signals can reach very small values, from milliseconds to macro seconds.

If the body moves in space, then it has a so-called shift in the frequency of the carrier signal f_0 - Doppler shift, thanks to this, it becomes possible to calculate data on the angular and radial velocities of the body, which arise due to the change in the angular increment and distance to the object over time.

The main component of the calculation of the Doppler shift F_D is the radial velocity V_r of the movement of the observation body, based on this, the formula was formed:

$$F_D = -\left(\frac{2 \cdot V_r}{c}\right) \cdot f_0 = -\frac{2 \cdot V_r}{\lambda_0},$$ (2.17)

where V_r is the radial growth rate relative to the movement of the observation object, λ_0 is the wavelength.

Relative to the movement of the body to the radar, further or closer, there is a proportional dependence of repulsion and reception of the signal by the radar system, that is, the further the object has moved, the longer the signal will reach and vice versa, which is not taken into account in the case of a constant position of the body. It follows that the phase of the received signal will have different values, which will be directly proportional to the change in the frequency of the radio signal. Thanks to these basic characteristics, the radial velocity of the body is calculated.

According to the principle considered when calculating the delay of the signal time difference for determining the angular coordinates, the difference in Doppler shifts determines the rate of change in the angular position.

Radar has the best properties in terms of obtaining a large amount of information, such as speed, position, frequency, initial phase, spectrum. However, all this becomes known only after the reflection of the directed signals, which creates a great stress on the system due to the volume of processed information and requires high accuracy of the transmission line. Thus, there is a problem related to the cost price, which can reach incredible proportions.

2.10. The method of creating 3D models using a set of a certain number of images

Many modern programs, such as Photomodeler, 123D Autodesk or 3DsMAX, have recently started to use and develop the technology of building a model thanks to a set of images [14], since this method greatly simplifies the work of designers, constructors, etc. Below we will consider how and in what sequence this method works.

For understanding, the scheme of photographing the object from different viewpoints is shown (Fig. 2.10). Looking at the given scheme, it is necessary to make a clarification regarding the overlapping of adjacent photos, a pair of photos

must have an overlap of adjacent frames of no worse than 50%, because otherwise, gaps may appear on the final 3D model. Looking at figure 2.13, adjacent pictures can be considered photo sets 1,2,3 or others.
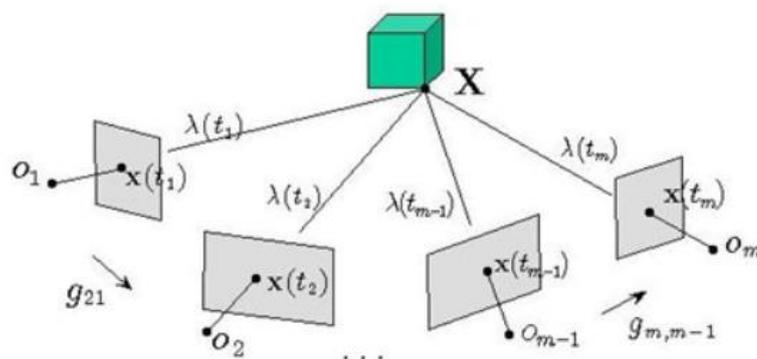
Fig. 2.13. Scheme of photographing the surveillance body

After taking a number of consecutive frames (Fig. 2.14), photo processing begins, which means finding the corresponding common points of the body in different frames, determining the oriented position of each picture relative to the initial, focal length and camera parameters, and based on the obtained ratios, calculating a system of nonlinear equations.
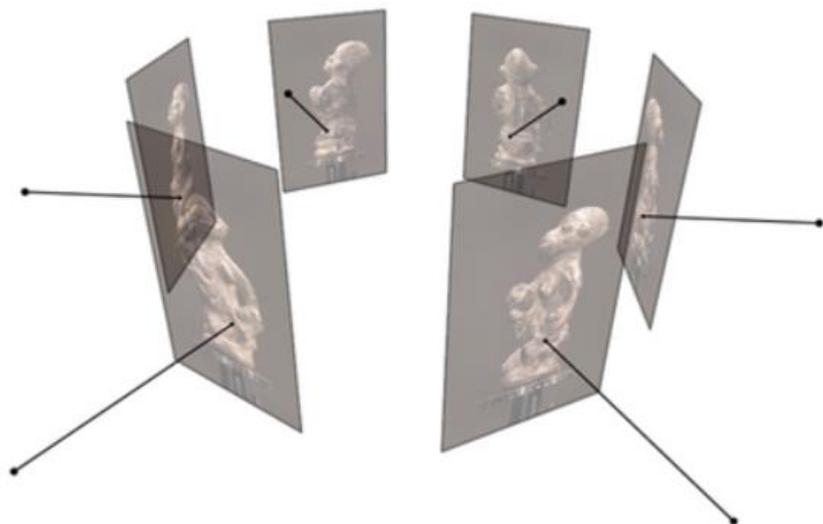
Fig. 2.14. Oriented position of each picture

After obtaining adjacent points for the corresponding pairs of images, a map of the depth of the image is formed, according to which the spatial coordinates of the position of the points are then calculated, which are displayed on the layout of the 3D model (Fig. 2.15).

Fig. 2.15. 3D model of the object

Delaunay triangulation builds a model from a set of points, which can then be textured or transformed through OpenGL.

Using this method of building 3D models is too expensive, because in order to capture one model, you need to create a mobile mechanism to move around, or use a fixed module, which requires installing a large number of cameras on the structure itself, or repositioning one camera between adjacent supports, with which means that a large amount of time or money is spent to create one model.

2.11. Using neural networks to create a depth map or volumetric model from an image

Over time, systems that use neural networks have gained momentum in development and use. Thus, when using artificial neural networks [15], the process of obtaining characteristic points and image processing is built into the very mechanism of operation. It is clear that appropriate use requires a set of training examples that can predict the desired outcome, i.e. training with a teacher, which in turn, from the perspective of stereo vision, requires obtaining a depth map in advance. For this, the cameras should be set up in advance, i.e., calibrated and data obtained from the internal and external parameters of the shooting device. To facilitate the process, it is possible to use the teaching method without a teacher, that is, without the use of advance calibration, but with a set of training examples.

An autoencoder is one of the methods of using a neural network without a teacher, which can reasonably find the influence and consequences of the input data that the system receives. For image processing, it is better to use the type of convolutional autoencoders that will be considered in this work.

If analyze the operation of a convolutional autoencoder, then an important characteristic is that during training in hidden layers, compressed conditions are created about the input data, which are then used for the image filtering process, for example, noise removal. After a series of experiments, it was found that after increasing the depth of the network and adding subdiscretization layers, provided that the number of convolution filters is reduced, the quality of image classification increases, and the computational load, on the contrary, decreases.

Research using a neural network [16] has shown that:

- The amount of computational load increases if the number of processing filters on the convolution layers is increased, and the accuracy of the result practically does not change;

- Selection of essential features in the image is carried out thanks to the recovery of the compressive transformation of the input image by the autoencoder;

- A small value of the depth of the network does not deteriorate the result shown by the autoencoder, thus, after conducting the experiment, it was proved that a small number of training parameters, provided that more than 80 iteration cycles are carried out, show a good accuracy result;

- The accuracy and color information of the image may be degraded, due to the fact that the dimension of the decoding and encoding boundaries is too low.

2.11.1. Using capsule neural networks to construct image depth

One of the advantages of capsule neural networks [17] is that, unlike convolutional neural networks, it is able to encode the orientation and position of an object in the image, after registering the presence of the object.
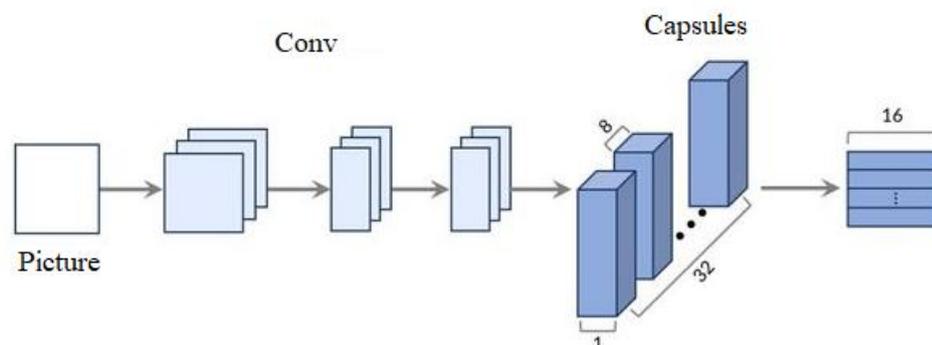
Fig. 2.16. The structure of the neural network of the capsule species

As the name implies, CapsNET (Fig. 2.16) for identifying patterns in the image consists of groups of neurons or so-called capsules. The process takes place step by step, the vector of information about the orientation and position of the frames goes to another layer, that is, the capsule, which processes the received input data and makes a prediction. Each capsule is responsible for a certain process and calculates the task regardless of the work of another capsule, that is, there is no connection between capsules of the same type. When dividing the output data from

the convolutional layer, the capsule itself is formed, thus the three-dimensional vector is divided so that the information about the pixel in the capsule is divided by three-dimensional coordinates. The orientation and position of the object or area in the image is fixed thanks to the state of the neurons of the capsule neural network.

The principle and operation of the capsule neural network is almost completely identical to the operation of a conventional convolutional network, but produces better accuracy results.

2.11.2. Using PlanetNET to construct image depth

PlanetNET works on the principle of learning without a teacher on a displacement map and without using two images of space. Dilated Residual Networks (DRN) are taken as the basis for creating Deep Neural Network (DNN) PlaNET, which will deepen the map after solving three sub-problems (Fig. 2.17):
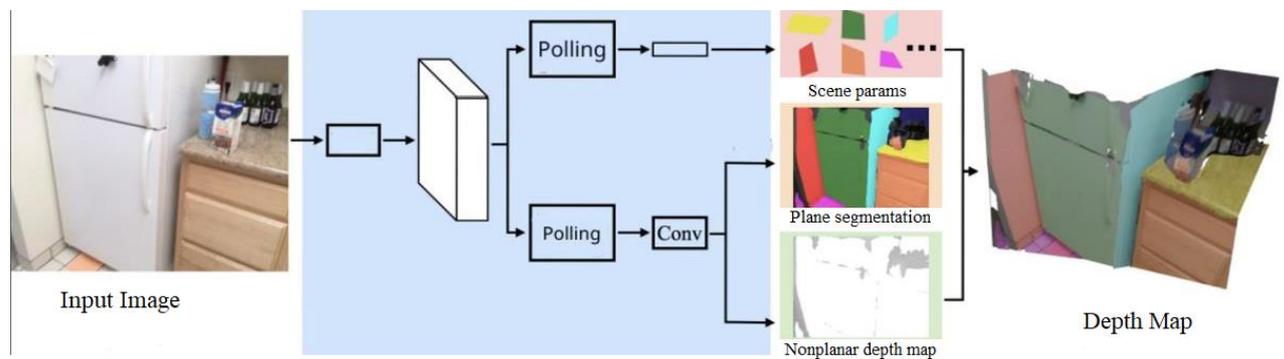
Fig. 2.17. Construction of a depth map using PlaneNet thanks to the prediction of parameters for one rgb picture: plane parameters, plane segmentation, non-planar depth map

- Finding the parameters of the plane means selecting and predicting the number of available planes K in the image, then selecting the levels of the plane K, which are characterized by three parameters P_i: line, normal and offset. The following formula was used to calculate the error:

$$L = \sum_{i=1}^{K} \min_{j \in |1,\widehat{K}|} ||\widehat{P_j} - P_i||, \qquad\qquad (2.18)$$

where P_i, K are the real parameters and the number of planes, and $\widehat{P_i}, \widehat{K}$ are the predicted corresponding values.

- Plane segmentation means identifying a group of pixels that belong to one plane of the body.

- To build a non-flat depth map, each pixel takes a depth value of 0 or 1, to create a one-channel map.

Using neural networks to solve the given problem is one of the best methods, but in order to create a quality product, the developer is required to create or have a good computing system or PC, another disadvantage is that the process of processing or training samples can be a large amount of time was spent, which is not expedient from a financial and time-consuming point of view.

CHAPTER 3. THEORETICAL INFORMATION ABOUT THE SOLUTION OF THE ASSIGNED PROBLEM

To solve the problem, an algorithm was chosen for development using stereo cameras [18], which allows viewing three-dimensional coordinates thanks to the use of two cameras that can find principle points, as shown in Fig. 3.1.
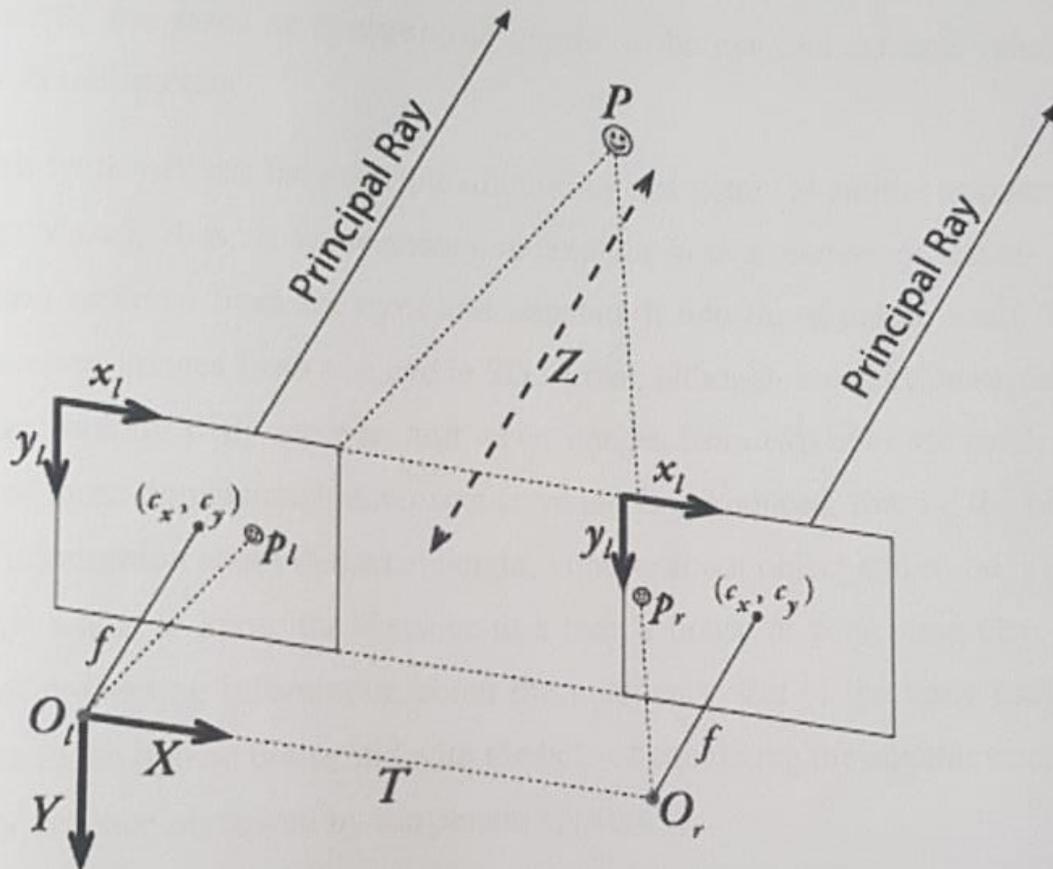
Fig. 3.1. Stereo coordinate system

The image shows that the optical centers of the lenses O_l, O_r are installed in parallel and must be at the same height, it is important to note that the focal length f of both devices must be the same or adjusted so that they coincide.

ACIC DEPARTMENT		NAU 22 1094 000 EN			
			N.	Page	Pages
Performed	O. S. Sych	IMAGE DEPTH DETECTION SYSTEM		57	160
Supervisor	M.P. Vasylenko				
Consultant			225 151		
S. controller	M.K. Filyashkin				
Dep. head	V.M.Sineglazov				

The focused object or body of interest to the observer is marked with the letter P, and the distance at which it is actually located is marked as Z, the unit of measurement of length can be whatever the developer wants, T is the distance between the optical centers, it is worth noting here that if it is planned to observe objects at long distances, then the distance between the centers should be increased, and if considering the vision corresponding to human, then the distance should be approximately the same as between the pupils of the eye, the average value of which is 55 millimeters.

This work will use the principle of building a structure as similar as possible to human vision, thus, it is necessary to explain how a person processes the information received from the eyes and translate it into the structure itself. The brain perceives images from one eye in 2D format, although it can evaluate depth even when looking with one eye, and when images from two eyes are used, the concept of three-dimensional stereoscopic vision is introduced, that is, the brain receives information about distance, height, volume about object and so on. Thus, if a person wants to know the distance to a tree, a house or something else, the process of processing information about the imaginary size of the body itself is carried out in the human brain, and with the help of equalizing the angular size, the imaginary distance perceived by the person appears.

According to the same principle, it is possible to configure any mechanism that can receive data from stereo vision. Below it will be considered how projective geometry is constructed.

3.1. Homogeneous coordinates and projective geometry

Projective geometry plays a big role in stereo vision. There are several types of solutions to projective geometry problems: analytical, geometric, and algebraic. The first two approaches refer to Euclidean geometry and have different properties, so that during the geometric approach, axioms about the geometric structure of the body are considered, and the analytical approach considers the coordinates of the

object itself. In general, the project will use an analytical method of solving projective geometry, an explanation of this method will be given below.

3.1.1. Points of the projective plane

As already understood, points on the projective plane have three coordinates $(x, y, w)^T$, that is, they represent a three-component vector, in contrast to the usual Euclidean plane, which is described by a pair of coordinates $(x, y)^T$. Let there be a certain point described by the coordinates $(x, y, w)^T$, then, provided that the factor a is not zero, the values $(ax, ay, aw)^T$ and $(x, y, w)^T$ refer to one point, while the zero vector $(0,0,0)^T$ is not considered and does not refer to any of the points. Provided that these properties of the points of the plane are preserved, homogeneous coordinates are considered.

A possible relationship between the Euclidean plane and the projective plane is as follows, for example, there is a Euclidean coordinate $\left(\frac{x}{w}, \frac{y}{w}\right)^T$ corresponding to a coordinate vector $(x, y, w)^T$, where $w \neq 0$, and if $w = 0$, then the coordinate vector is considered in the form $(x, y, 0)^T$, which implies finding a point at the so-called infinity. Based on this, the Euclidean plane is considered as a projective plane to which infinity points are added.

In order to make a transition between coordinate types, there is need to make certain mathematical transformations, for example, if there is need to go from Euclidean to uniform coordinates, then you need to add one to the initial position vector $(x, y)^T$ and go to the form $(x, y, 1)^T$. If it is necessary to make the inverse transition from homogeneous coordinates $(x, y, w)^T$, it is enough to divide the vector by its last value, that is, w, then the vector will acquire the Euclidean form of coordinates $\left(\frac{x}{w}, \frac{y}{w}\right)^T$.

3.1.2. Mapping of the straight lines on the projective plane

Just as the coordinates on the projective plane are described, the line on the plane can be described using a three-component vector by the expression $l = (a, b, c)^T$. If the equation of a straight line is described with precision up to a non-zero factor, then it takes the form $l^T \cdot x = 0$.

If the condition $a^2 + b^2 \neq 0$ is fulfilled, then the equation of the line will be considered in the form $a \cdot x + b \cdot y + c = 0$, and the line is at infinity, provided that the vector acquires the values $(0,0,w\,)$.

3.1.3. Three-dimensional projective space

According to the same principle, as the projective plane is determined, the three-dimensional vector of the projected space changes with the addition of one more parameter, and now the four-component vector has the form $(x, y, z, w)^T$. Similarly, the change occurs for any non-zero number a, then the coordinates are written in the form $(ax, ay, az, aw)^T$ and must correspond to the same point with coordinates $(x, y, z, w)^T$.

In the same way as was discussed above, the dependence between the points of the three-dimensional projective space and the Euclidean space is established, provided that $w \neq 0$ in the vector of uniform coordinates $(x, y, z, w)^T$, and the point with coordinates $(x, y, z\,, 0)^T$ is assumed to be at infinity.

3.1.4. Conversion to projection

An important aspect of this method is the projective transformation, which performs an inverse transformation on the projective space or plane, i.e. transforms line segments into lines. For this, it is customary to derive a non-degenerate projective matrix H in coordinates, which is a transformation of the coordinates of the vector x to the form x' - according to the formula $x' = H \cdot x$.

3.2. Projective camera model

The cameras used in the modern world are well suited for models that describe the concept of a projective plane. When considering the camera model, it is necessary to start from the center of the camera, the main axis is considered to be the line directed from the center of the camera to the central point of the observation plane, the image plane is considered to be the area intercepted by the camera, which will then be projected into a cloud of points representing the coordinates of the points of each of pixel in the image. Any point X of the observed region on the plane is projected onto the image plane by the corresponding point x located on the segment CX, which connects the starting point X with the center of the camera C (Fig. 2.3).
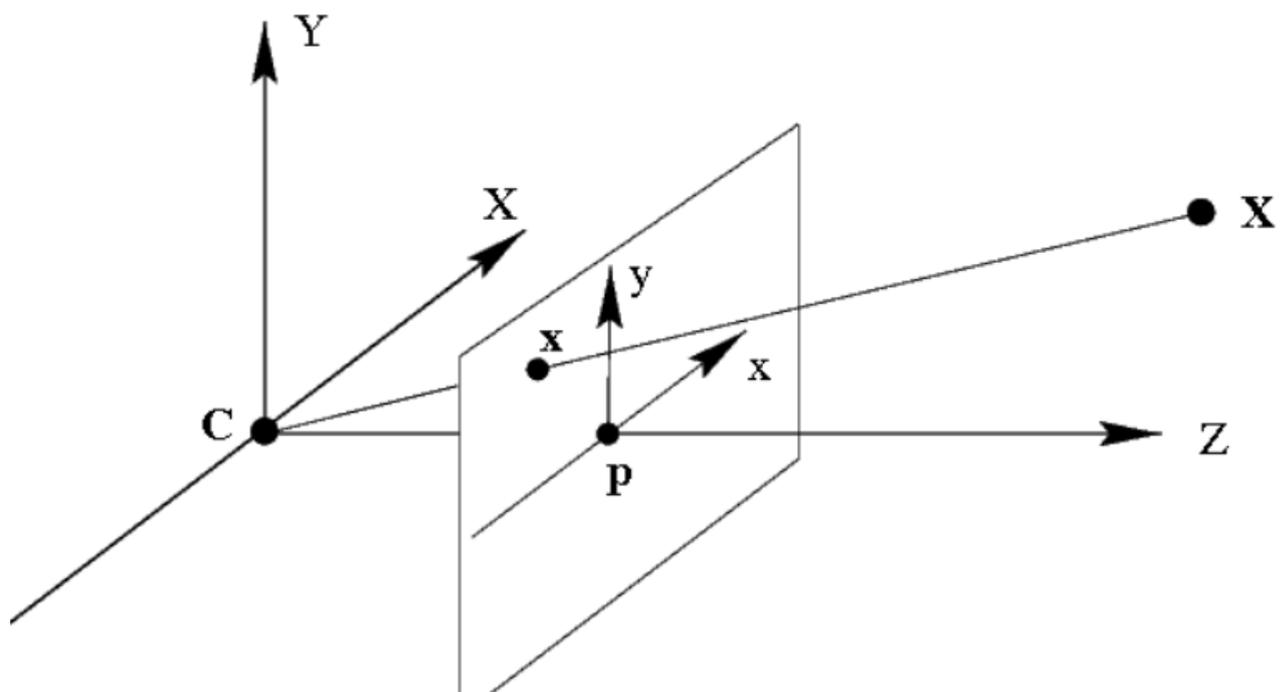
Fig. 3.2. Camera model. Z is the main and main axis directed from the center of the camera, C is the center point of the camera. The point x is projected onto the image plane, from the starting point X of the three-dimensional space.

In uniform coordinates, the mathematical model of writing the design formula has the form:

$$x = P \cdot X \qquad (3.1)$$

Where P is the matrix of the camera, which has a size of 3×4, x are the points of the plane of uniform coordinates, X are the points of the space of uniform coordinates.

To express the matrix P, it is necessary to solve the formula:

$$P = K \cdot R \cdot [I \,|{-c}] = K \cdot [R|t] \qquad (3.2)$$

Where R is an orthogonal matrix that includes data on the tilt of the camera in global coordinates, that is, relative to each of the x, y, z axes, the dimension of this matrix is 3×3, I is a matrix that is unitary with a size of 3×3, K is a triangular matrix that stores the internal parameters of the camera and has a size of 3×3, c is the coordinates that correspond to the center of the camera, $t = -R \cdot c$.

The results of projecting points will not be modified, because the camera matrix is predetermined, which corresponds to the precision of the specified non-zero multiplier and corresponds to the ratio $x = P \cdot X$. The multiplier is chosen in such a way that the above conditions are properly fulfilled.

The coordinate axes on the camera plane will have the same scale in the case when the CZ axis is equally directed with the central axis of the camera, and the center of the camera, in turn, lies at the origin and has zero coordinates, in which case the camera matrix will be calculated as $P = K \cdot [I|0]$, where

$$K = \begin{pmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{pmatrix}. \qquad (3.3)$$

In most modern cameras, the image center can have non-zero coordinates, and the pixels displayed after taking a photo or video camera are slightly different from square pixels. Taking into account all the above, the matrix of internal parameters will look like this:

$$K = \begin{pmatrix} a_x & 0 & x_0 \\ 0 & a_y & y_0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{3.4}$$

The coefficients α_x and α_y are considered to be the focal lengths f along the x and y axes.

Distortion of the image that occurs due to the unclear adjustment of the optics, the data obtained from the errors along the x and y axes are written in the form of formulas:

$$\begin{cases} x'' = x' \cdot (1 + k_1 \cdot r^2 + k_2 \cdot r^4 + k_3 \cdot r^6) + \\ \quad + 2 \cdot p_1 \cdot x' \cdot y' + p_2 \cdot (r^2 + 2 \cdot x'^2) \\ y'' = y' \cdot (1 + k_1 \cdot r^2 + k_2 \cdot r^4 + k_3 \cdot r^6) + \\ \quad + p_1 \cdot (r^2 + 2 \cdot y'^2) + 2 \cdot p_2 \cdot x' \cdot y' \end{cases} \tag{3.5}$$

Where k_1, k_2, p_1, p_2, k_3 are the parameters of the optical system describing the distortion coefficients; x'', y'' — distorted coordinates of the point according to the center of the image, provided that square pixels are used; (x', y') — projection coordinates of points according to the center of the image, provided that square pixels are used and without distortions; $r^2 = x'^2 + y'^2$.

One of the features of distortion is that it does not depend on the distance at which the object of interest is located, but depends only on the projection of the

object's pixels, that is, the coordinates of the points. In order to avoid distortion, the original image is transformed by means of image filtering from the camera. Only if the focal length remains constant will the transformation be the same for the images taken by the camera, i.e. the matrix of internal parameters of the camera remains the same.

It is customary to say that a camera is considered calibrated only when the distortion factors and internal parameters of the camera are known.

3.3. The working principle of stereo vision or the relationship between two cameras

An important role in the construction of the depth map is played by the matrix of reflections of two cameras, which determine three-dimensional coordinates, namely, the process of calibration of stereo vision.

Suppose that there are matrices P and P' of two cameras in some coordinate system. In the case when the directions of the two central rays of the cameras do not coincide, it is possible to determine the coordinates of the points of the three-dimensional plane, and the cameras are considered calibrated.

Camera matrices are described according to the coordinate system and written in the form:

$$P = K \cdot [I|0], \tag{3.6}$$

$$P' = K' \cdot [R'|t']. \tag{3.7}$$

If the Z axis is directed along the optical axis, which coincides with the center of one of the cameras and the origin of coordinates. Highlighting the key points of the image occurs thanks to repeated shooting of the observed plane on which it is easy to determine the position of objects relative to others in space, this process is

called camera calibration. After receiving the input data, the mathematical processing of information about the position of template points in space relative to the camera matrix, which are connected by the coordinates of the projection itself, begins.

3.3.1. Epipolar geometry

In order to understand the description of the method of calculating the three-dimensional coordinates of points, it is worth starting with the explanation of the geometric properties on the plane, which are closely related to the position of both cameras relative to the projection of the position of points in three-dimensional space.

Let's assume the existence of two primary components, as in Fig. 3.3. The first chamber's origin is C, while the second chamber's center is C'. The left camera's image plane and the right camera's image plane are each projected with a point in space, X, at x and x', respectively. The ray xX is the pre-image of the point x in the left camera's view. The epipolar line, also known as the projection plane, is a straight line that this beam is projected onto in the second chamber. The epipolar line l' is inextricably located on the picture of the point X on the image plane of the second camera.

As a result, every point on the left camera's picture correlates to an epipolar line on the right camera's image, or vice versa. The pair for x in the right camera's view can only be located on the appropriate epipolar line in this situation. The epipolar line l' on the left picture corresponds to each point x' on the right image.

To find stereopairs and determine if a pair of points can be a stereopair, epipolar geometry is utilized (i.e., a projection of some point in space). The coordinate notation for the epipolar geometry is relatively straightforward. Assume that there are two calibrated cameras, and that x and x' represent the homogeneous

coordinates of a point on the images of the first camera and the second camera, respectively.
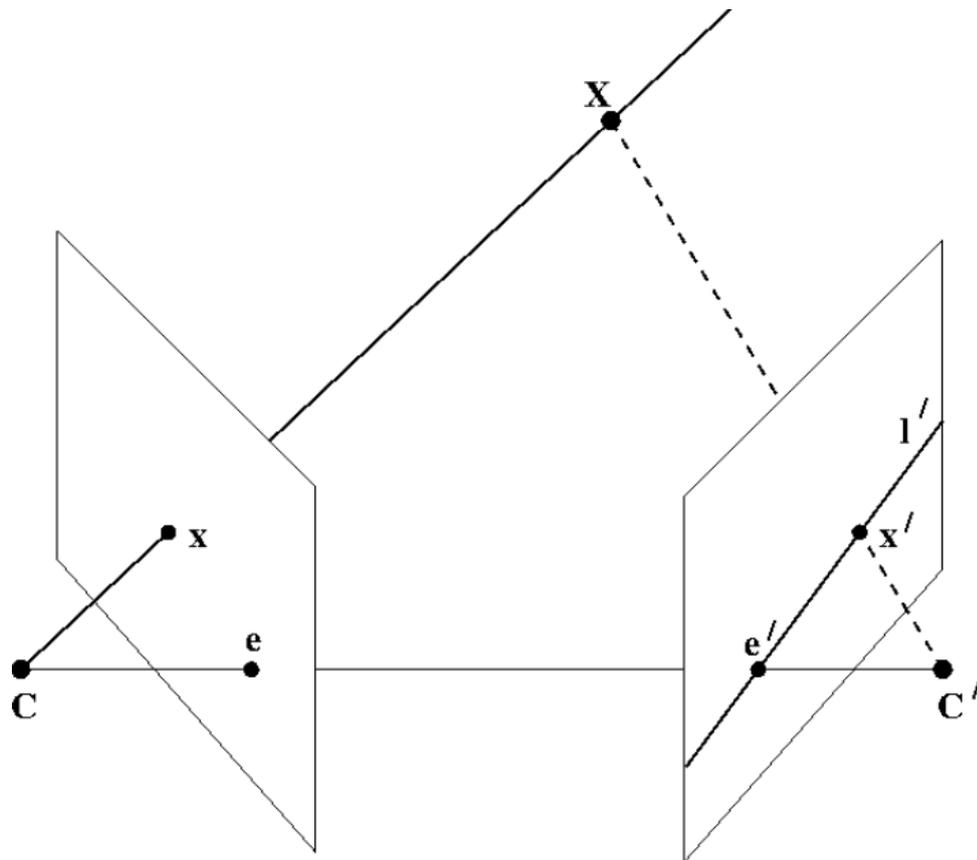
Fig. 3.3 Epipolar geometry

A pair of points x, x' is a stereopair in a 3×3 matrix F when and only when

$$x'^T \cdot F \cdot x = 0. \tag{3.8}$$

The fundamental matrix is designated as F. It simply depends on the matrices of the initial P and P', has a rank of two, and is defined up to a non-zero factor.

When the camera matrices appear as follows:

$$P = K \cdot [I|0], \tag{3.9}$$

$$P' = K' \cdot [R|t], \qquad\qquad (3.10)$$

and the formula may be used to determine the basic matrix:

$$F = K'^{(-1)^T} \cdot R \cdot K^T \cdot [K \cdot R^T \cdot t] \cdot x, \qquad\qquad (3.11)$$

where the sign $[e]_X$ is obtained in terms of the vector e which contain epipolar coordinates:

$$[e]_X = \begin{bmatrix} 0 & -e_z & e_y \\ e_z & 0 & -e_x \\ -e_y & e_x & 0 \end{bmatrix}. \qquad\qquad (3.12)$$

The epipolar line equations are calculated using the basic matrix. The equation of the epipolar line itself is $l'^T \cdot x' = 0$, and the vector defining the epipolar line at a point x will take the form $l' = F \cdot x$. The vector denoting the epipolar line will consequently appear as $l = F^T \cdot x'$ for the point x'.

The essential matrix seems to be another option in addition to the basic matrix:

$$E = K'^T \cdot F \cdot K. \qquad\qquad (3.13)$$

The essential matrix and the fundamental matrix will match if the internal parameter matrices are both singular. The important matrix is utilized in situations

where it is necessary to ascertain the movement of the camera since it may be used to recover the position and rotation of the second camera in relation to the first.

3.3.2. Triangulation of pixels or points in an image

Let's really go on to discussing how to calculate a point's three-dimensional coordinates from its projection coordinates. Triangulation is the name given to this phenomenon in the literature.

Two calibrated cameras with the matrices P_1 and P_2 should be present. The projections of a certain point in the space X are given by the homogeneous coordinates x_1 and x_2. The system of equations that results may then be written as follows:

$$\begin{cases} x_1 = P_1 \cdot X \\ x_2 = P_2 \cdot X \end{cases}.$$

(3.14)

In real life, this system is solved using the following method. Eliminate linearly dependent equations, vectorially multiply the first equation by x_1, the second by x_2, and get the system to the condition $A \cdot X = 0$, where A has a size of 4×4. If the vector X is the homogeneous coordinates of the point, you may either continue from there or set its last component equal to 1 and solve the ensuing system of three equations with three unknowns. An alternate approach is to use any non-zero solution of the equation $A \cdot X = 0$, such as the singular vector computed as the matrix A's lowest singular value.

3.4. Depth map construction

The difference between each pixel's distance from the camera and its color is saved in a depth map [19], which is an image. A depth map can be created from

a stereo pair of pictures or by employing a specialized depth camera (the Kinect sensor is an example of such a camera).

A stereopair-based depth map may be created using a fairly straightforward concept. A search is done for each point on one picture to find its associated point on the second image. Additionally, you may triangulate two related locations to find the three-dimensional coordinates of their pre-image. The depth is determined by multiplying the three-dimensional coordinates of the pre-image by the separation from the camera plane.

On the epipolar line, one must look for the paired point. In order to make the search process simpler, the pictures are positioned so that all of the epipolar lines are parallel to the image's sides (usually horizontal). Additionally, the pictures are positioned so that for each point, the paired point corresponding to it must be found on the same line on the image from the second camera. For example, if a point has the coordinates (x_0, y_0), the epipolar line corresponding to it is provided by the equation $x = x_0$. Rectification is the name given to this alignment of the images. In most cases, rectification entails remapping the picture in order to remove distortions.

When the photos have been corrected, the matching pairs of points are looked for. The simplest method is shown in Figure 3.4 and goes like this. A search is done for a pixel in the right picture for each pixel in the left image with the parameters (x_0, y_0). The pixel in the right image is believed to have coordinates $(x_0 - d, y_0)$, where d is a number known as disparity or displacement (disparity). Calculating the maximum of the response function—which may, for instance, be the correlation of pixel neighborhoods—is done in order to find the associated pixel. A disparity map is the end product.
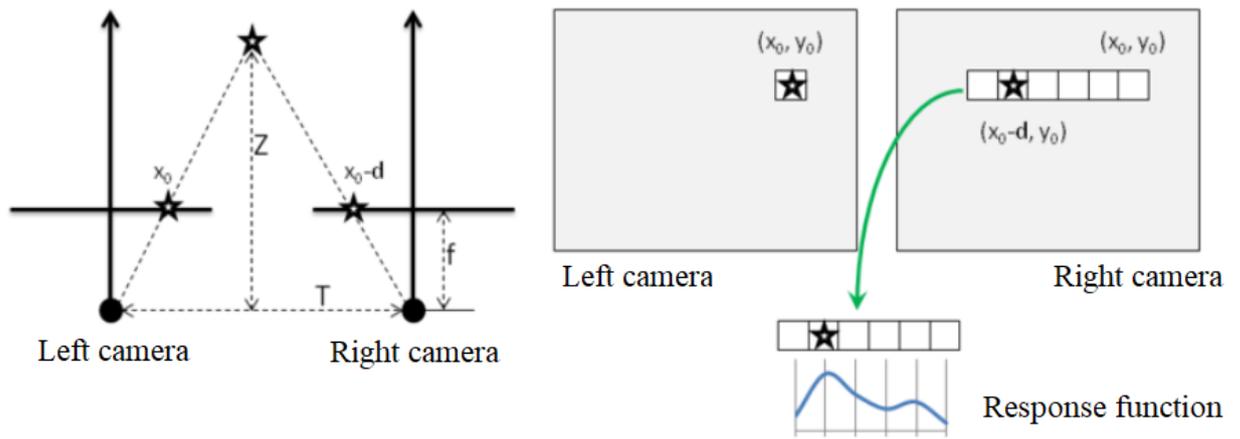
Fig. 3.4. Basic principle of depth map construction

In reality, the relationship between the depth values and pixel offset is inverse. The link between disparity and depth may be stated as follows using the notation from the left side of figure 3.4:

$$\frac{T-d}{Z-f} = \frac{T}{Z} \rightarrow Z = \frac{f \cdot T}{d}. \tag{3.15}$$

The resolution of stereo vision systems that operate on this principle is better at close distances and worse at vast distances because of the inverse connection between depth and displacement.

3.5. Image filtering.

Images are subjected to a variety of interference during conversion and transmission utilizing wireless networks, which in certain situations can lower the quality of the image and result in the loss of its fragments. Solving the issue of image reconstruction for filtering pictures collected with the use of photo or video cameras is becoming more and more critical as digital communication systems proliferate. In actuality, noise that develops throughout the stages of picture

synthesis and transmission over the communication channel frequently distorts images.

The three subfields of computer graphics are visualization, image processing, and pattern recognition. The act of reproducing an image from a description is known as visualization (model). Extraction of a semantic description of the things shown in a picture is the primary goal of image recognition. Image transformation is image processing (filtering). Automatic image processing systems are extensively used in practice thanks to the quick growth of computer hardware and information technologies.

Enhancing image quality is the major objective of these systems. One of the most important and well-known issues in image processing is noise suppression. Gaussian noise, impulsive noise, and their symbioses are the two most prevalent kinds of noise.

The enhancement (recovery, restoration) of an image in accordance with a general criterion is one type of request for image processing, while another type is a particular transformation that accommodates an evolving image. Image processing may be a preliminary step before storing the image in the latter scenario (for example, selecting the outline of the object). The way an image is produced—whether it is created by a computer graphics system by digitizing a color or black-and-white photograph or video—has a significant impact on the image processing techniques used. If the image was created by digitization in that situation, noise is typically present.

In medicine, noise reduction can be used to sharpen the picture on X-ray imaging, as a pre-processing for subsequent detection, etc. Noise reduction is most frequently employed to enhance visual perception, but it can also be used for some particular applications. In picture compression, noise reduction is crucial. The quality of the compressed image may suffer if there is too much noise present during compression that may be misinterpreted for visual detail.

1) Inadequate camera, scanner, etc., for taking pictures;

2) unfavorable lighting circumstances;

3) interference during transmission across analog channels; interference from sources of electromagnetic fields; inbuilt noise of active components (amplifiers) of the wire.

Types of noises [20]:

Accordingly, noises also appear in different forms. Additive models of Gauss and impulse noise are most suitable for use in practical tasks. Additive Gaussian noise is characterized by the fact that pixel values from a normal distribution with a mean value equal to zero are added to each image. Such noise is usually introduced at the digital display stage. Impulse noise is characterized by the replacement of some image pixels with values of some fixed or random value. Such a noise model is associated, for example, with image transmission errors. Noise removal methods.

Algorithms for noise reduction often focus on decreasing a particular kind of noise. There are yet no filters that can be used universally to find and eliminate all noise. However, most algorithms concentrate on suppressing this kind of noise as many noises can be adequately represented by a Gaussian white noise model.

The same widely used noise-cancelling techniques [21]:

- smoothing filter;
- Wiener filter;
- median filter;
- rating filter.

There are both linear and non-linear filters used to suppress Gaussian noise. A function (filter kernel) with a specific effective sign on the grid defines a linear filter. Discrete convolution processes are used to carry out the filtering itself

(weighted summation). The intensity value at each point is averaged over a smoothing mask in linear smoothing filtering. In the first scenario, the center point's intensity value is equal to the average of the intensities around. In other instances, the coefficients' weighted average value is provided. It is possible to improve image processing outcomes, particularly filtering, by using a Wiener filter. Its application is connected to the idea of picture stationarity. Since picture edges interfere with stationarity, Wiener filtering isn't ideal strictly speaking.

The influence on the level is substantial for picture sizes that greatly exceed the correlation interval and the image. Technically, discrete Fourier transforms in the frequency domain are used to accomplish Wiener filtering.

However, many issues of practical significance cannot be satisfactorily solved when linear filtering techniques are used. It is important to remember that non-linearity permeates even the processes of information transmission, coding, and perception, including information sensors, communication channels, and the human visual system. In order to increase the spectrum of problems that digital image processing tools can handle and get around the drawbacks of linear filtering techniques, nonlinear digital filtering approaches are actively being used.

The theory of nonlinear filtering [22] is extremely difficult to construct in a single step, in contrast to the theory of linear filtering. Each of the aforementioned classes has unique benefits and various application areas. In photos that have been affected by impulsive noise, for instance, median filters are recognized to produce the greatest results in retaining color differences, borders, and local brightness peaks.

The median filter is a non-linear decolorization procedure, in contrast to the smoothing filter. The sliding window w used by the median filter covers an odd number of samples. The median of all the picture components in the window takes the place of the media sample. The average value of the generated rows of the original array is the median of a discrete array.

The median filter is used to reduce additive and impulsive noise in the picture, just as the smoothing filter. The median filter's ability to maintain variations in brightness sets it apart from smoothing filters (contours). The median filter will perform better than the ideal linear filter if the brightness gradient is strong relative to the variance of the additive white noise. When dealing with impulsive noise, the median filter is particularly useful.

The ranking filter alters the picture using a mask [23], much like the smoothing filter. The central pixel may or may not be covered by the mask. Certain moments of this series, such the mean intensity or variance, may be determined since the values of the mask's component parts are placed in an ordered series in ascending (or descending) order. The median of the resulting series and the weighted sum of the intensities of the center pixel serve as the output value of the filter that replaces the central sample. The pixel statistics in the filter window have a particular connection with the coefficients.

The techniques mentioned above can progressively enhance the quality of distorted photos by smoothing out their edges. Component and hybrid filters perform the poorest when compared to linear and median filters. This relates to the computation of subfilter output (linear and median filters). The best method for removing Gaussian noise from an image is linear filtering, which obliterates the image's features. Using a non-linear filter based on rank statistics to eliminate impulsive noise while maintaining all visual variations is the simplest method. It is easy to employ algorithms that take into consideration the features of this noise if complicated noise has to be removed from the image (complex filters, hybrid filters, adaptive filters). It is based on both linear and non-linear filtering methods.

Three categories may be used to categorize computer graphics: visualization, image processing, and pattern recognition. Visualization is the recreation of an image from a description (model). Getting a semantic description of the things shown in a picture is the basic goal of image recognition. Image transformation is

image processing (filtering). Automatic image processing systems are widely used in daily life as a result of the advancement of computers and information technology.

CHAPTER 4. AN EXPERIENTIAL SOLUTION TO THE PROBLEM OF IMAGE DEPTH ESTIMATION

Numerous gradual modifications and adaptations for the study setting were made in order to resolve the problem of identifying the depth map.

The discussion of each of these things will follow.

4.1. Choosing a camera

Due to the need to create a more affordable and effective version of the project, two UTM Webcam (SJ-922-1080) web cameras were selected based on their features and cost (Fig. 4.1).

Fig. 4.1. The initial view of the camera UTM Webcam (SJ-922-1080)

The list of qualities of UTM webcam are listed in Table 4.1.

ACIC DEPARTMENT			NAU 22 1094 000 EN			
				N.	Page	Pages
Performed	O. S. Sych		IMAGE DEPTH DETECTION SYSTEM		76	160
Supervisor	M.P. Vasylenko					
Consultant				225 151		
S. controller	M.K. Filyashkin					
Dep. head	V.M.Sineglazov					

Table 4.1. The characteristics of the camera

Characteristic	Availability
Microphone	High-sensitivity microphone built-in FullHD with a wide angle
Sensors	CMOS sensor
Resolution	FullHD (1920x1080) video resolution
Focus system	Autofocus for focusing
Shooting frequency	20 frames per second
Overview	90-degree field of view
Fixings	Desktop mount, clothespin
System compatibility	Windows 10, Windows 7, Windows 8, Linux 2.6.21, Android 5.0, Windows Vista, Windows XP, and macOS

A collapsible design and the option to choose between three different program resolutions (MJPG 640×360, MJPG 1280×720 and MJPG 1920×1080) were also utilised throughout operation.

4.2. Preparation of the experimental setup

The simplest but most rigid connection was made between a board with a size of $115 \times 600 \times 20$ mm and a shaped wooden corner located in the center of the board; two cameras were fixed from above with an aluminum profile $300 \times 25 \times 3$ mm. This was done because it was necessary to create a stable structure in order to get adequate results during the study and because the cameras needed to be parallel and at the same height. The installation thus resembles the one in Fig. 4.2.

Fig. 4.2 Installation type

Almost every contemporary phone comes with a compass, which may be used to verify the plane of the stand to ensure it is horizontal after usage. It is also important to note that adding stable axes of movement, such as rollers for technical structures, is necessary if the experiment is being conducted in a moving plane. This is because shifting poses a risk of displacement with respect to the horizon axis, which can have a significant impact on the work's effectiveness.

4.3. Setting up cameras for software usage

The cameras must be positioned parallel to the structure's X and Y axes in order to get an appropriate result. For this, much as the entire project, the code (Appendix A) was created based on the MATLAB program. Algorithm for showing parallel axes from video cameras step by step is shown on Fig. 4.3.
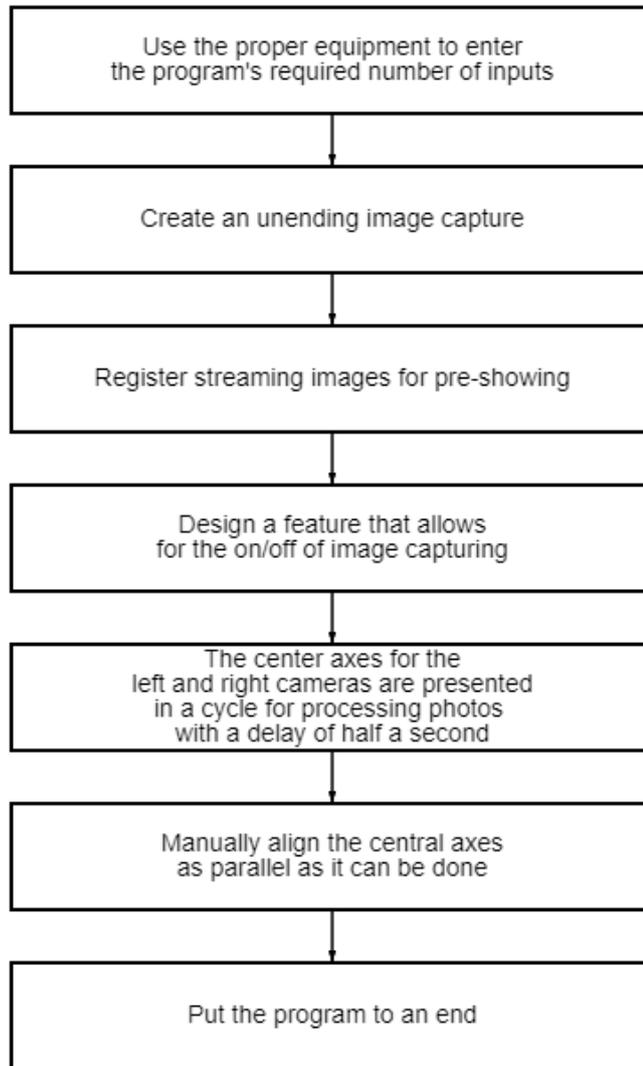
Fig. 4.3. Algorithm for showing parallel axes from video cameras

An illustration of this may be seen in Fig. 4.4, where the cameras must be manually adjusted and the axes are depicted with red lines.
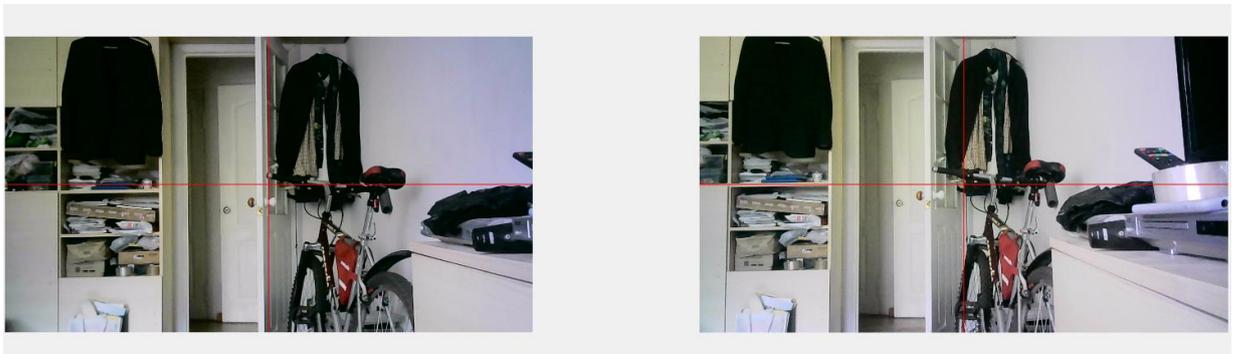
Fig. 4.4. Central axes on two cameras

4.4. Calibration process for two cameras

Evaluation of camera parameters is the process of calibration. In order to accurately determine the relationship between a 3D point in the real world and the corresponding 2D image, the projection (each pixel) in the image captured by the calibrated camera, the future work program will have access to all the information (specifications or design factors) about the camera.

For this, typically two types of parameters are required:

- The camera/lens system's internal parameters.
- External characteristics related to the orientation (rotation and displacement) of the camera with respect to some global coordinate system, such as the focal length, optical center, and radial distortion of the lens.

4.4.1. Camera calibration methods

Following is a list of the primary techniques for calibrating a camera:

1) Template Calibration: The ideal technique to calibrate is to take several photographs of an item or template with known sizes from various angles while having complete control over the imaging process. This is precisely the case with the checkerboard approach. Known-size circular designs can be used as an alternative to a checkerboard pattern;

2) Primitives in geometry. Other geometric primitives, such straight lines and junction points, can occasionally be found in the picture and utilized for calibration;

3) Deep Machine Learning may be used to gather data for camera calibration even when the options for manipulating picture parameters are highly constrained (for instance, a single photograph of a scene).

4.4.2. Algorithm for camera calibration

A template calibration approach was used in this work.

The block diagram below explains the calibrating procedure (Fig. 4.5).
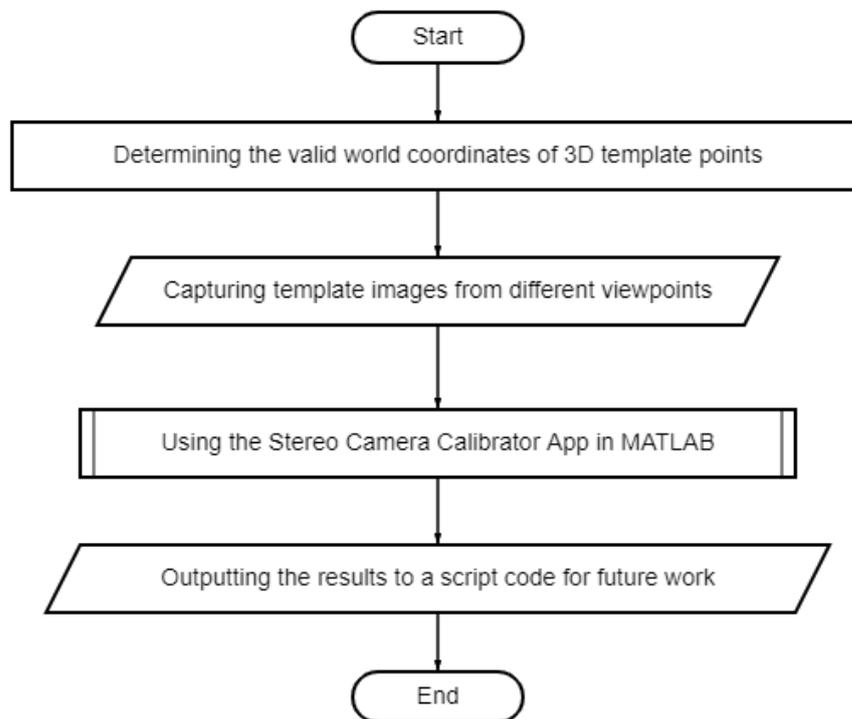
Fig. 4.5. Block diagram for camera calibration

1) Figuring out the correct world coordinates for 3D template points.

The checkerboard cells' corners are 3D points. As the starting point for the global coordinate system, you can choose any corner of the board. The Z_w axis is parallel to the board and perpendicular to the X_w and Y_w axes. As a result, $Z_w = 0$ and all chessboard points are on the XY plane.

A checkerboard with given dimensions is shot in various orientations to create 3D points (Fig. 4.6). The chessboard is connected to the global coordinate. All corner points are on a plane, therefore the observer can select Z_w at random for any point that is equal to 0. The chessboard's points are evenly spaced apart, therefore it is simple to calculate the coordinates (X_w and Y_w) of each 3D point by choosing one point as the origin (0, 0) and measuring the other points in relation to it.
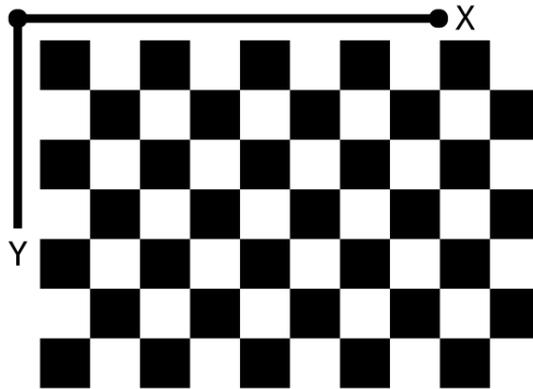
Fig. 4.6 World coordinate system: A pin-attached checkerboard pattern represents world coordinates

The graphic makes it simple to identify and locate the checkerboard squares. The checkerboard square corners are also perfect for localizing them because they have steep slopes in both directions. The junction of the checkerboard lines is also where the corners are located. The accuracy of the search for the corners of staggered squares is determined by all of these variables.

2) Take pictures of the template from various angles.

A code (Appendix B) was constructed to carry out this algorithmic stage, which involved reading the pictures from the left and right cameras and then creating a cycle with the goal of obtaining 50 photos as the final condition. A time delay, which can be adjusted to meet your needs and requirements, was introduced to allow you to adjust the position when capturing a photograph. Consequently, two pictures were shot at once (Fig. 4.7).
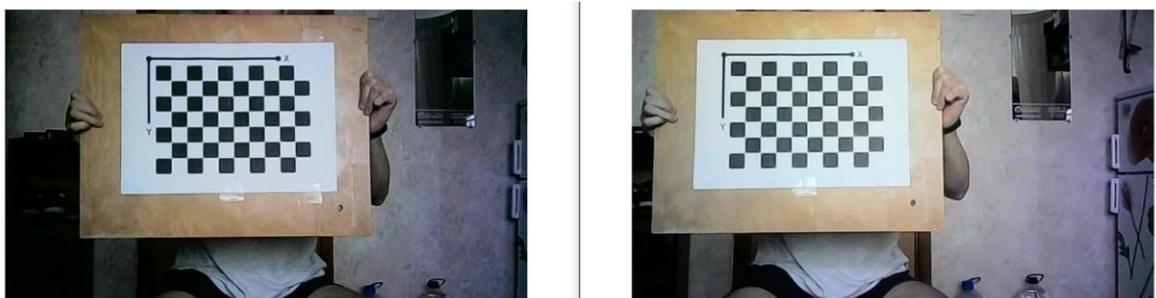
Fig. 4.7. The left and right photos being captured simultaneously

3) Using the app for calibrating stereo cameras.

MATLAB software component calibrates stereo cameras so it may be used to capture 3D sceneries, collect depth information from images, and other tasks.

The user must first input photos from the left and right cameras, as well as the precise dimensions of the checkerboard squares, to the interface before beginning work. The distance is then measured with millimeter-level accuracy. The examples that work may then be examined, and the radial distortion by 2 and 3 coefficients computation, as well as the skew and tangential distortion parameters, can be included. In the context of the work that was done, the skew and tangential distortion parameters were selected as the two coefficients. Figures 4.8, 4.9 and 4.10 show the outcome; during processing, certain frames were eliminated because their effects had a detrimental impact on the average inaccuracy per pixel.
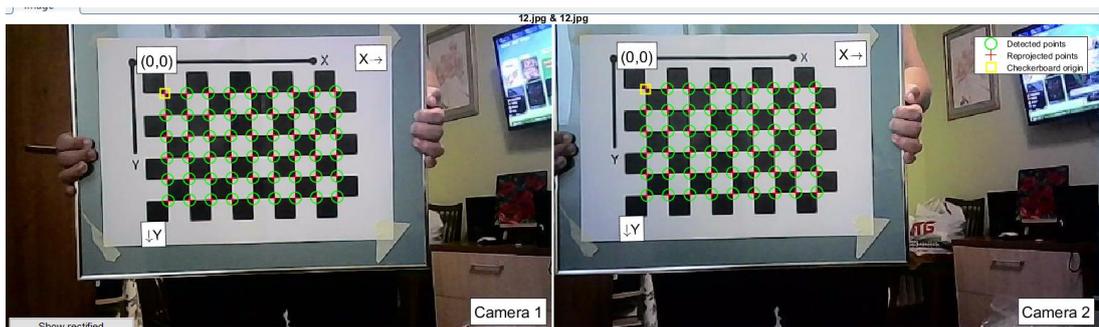
Fig. 4.8. The coincidence of the beginning points and axes of the template
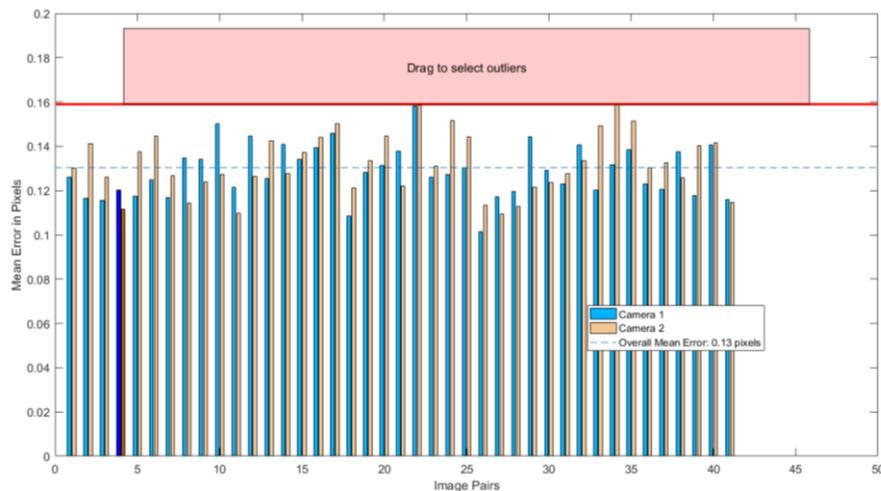
Fig. 4.9. Average error per pixel in different pairs of pictures
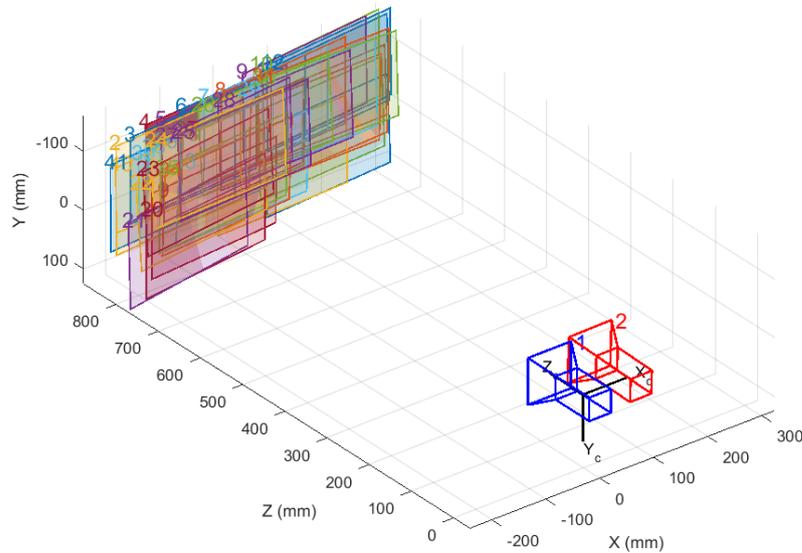
Fig. 4.10. Location of the cameras in space and according to each other

4) Outputting the data that was received to the script code.

In order to avoid having to do it manually every time the "Load" command is written in the program code, which loads all the used data, all incoming data will be added to the script code, which must be stored beforehand and manually invoked (Fig. 4.11).

stereoParams ✕	
1x1 stereoParameters	
Property ▲	Value
CameraParameters1	1x1 cameraParameters
CameraParameters2	1x1 cameraParameters
RotationOfCamera2	[0.9999 0.0042 -0.0143;-0.0048 0.9989 -0.0460;0.0141 0.0460 0.9988]
TranslationOfCamera2	[-79.1859 2.7748 -3.7986]
FundamentalMatrix	[-5.0030e-08 7.6865e-06 0.0028;-1.0308e-05 -7.5188e-06 0.1190;-0.0029 -0.1147 0.6538]
EssentialMatrix	[-0.0237 3.6670 2.9465;-4.9286 -3.6227 79.0406;-3.1067 -79.0877 -3.6852]
MeanReprojectionError	0.1304
NumPatterns	41
WorldPoints	54x2 double
WorldUnits	'mm'

Fig. 4.11. Script code's generated stereo parameters

4.5. Creation of the code for the depth map calculation

It was essential to determine and create an operation algorithm (Fig. 4.12) for at least two photographs deleted from the installation before generating the code

for processing streaming, which was completed, and then add the capability to see live objects. How this operation was carried out is discussed in detail below.
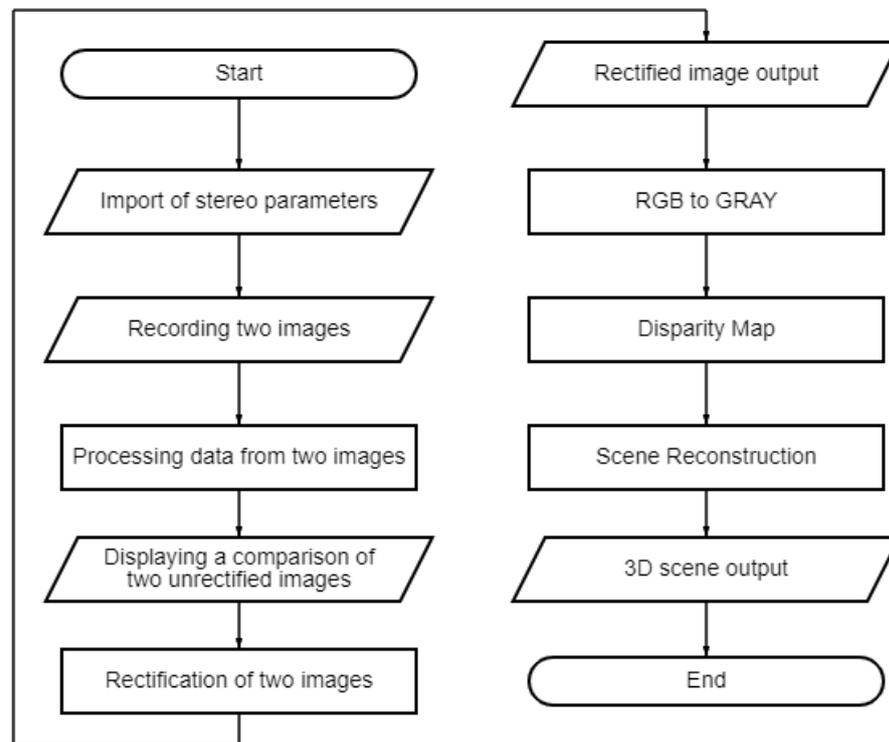
Fig. 4.12. The program's algorithm

1) Stereo parameters are one of the most crucial components of this job, as previously said, and their direct entry into the code enables the development of a software tailored to the user's demands. The position of the cameras in space is presented on the screen during import, as illustrated in Fig. 4.10, and all data is transmitted.

2) The RGB image is then read in uint8 format (it represents an integer from 0 to 255 and requires 1 byte); the size of the image depends on the shooting resolution chosen; to obtain the desired result, a 640x360-format image will suffice. The Gaussian filter, which calculates the transformation that is applied to each pixel of the two pictures using the normal distribution, makes it possible to smooth the image in order to draw conclusions that are more accurate. Each component of the picture has a corresponding weighting factor value. The weighting function is made

up of the total of all the weighting elements. The size of the window must be odd in order to identify the primary piece clearly. A 3×3 convolutional window was consequently selected. The convolution kernel enables the augmentation or dilution of picture elements. The weighting function is unaffected by the movement of the matrix across the picture.

3) The result of two unrectified pictures (Fig. 4.13) demonstrates a Y discrepancy that calls for correction in the next steps.

Fig. 4.13. Unrectified Stereo Image

This is resolved, as shown in the image, by superimposing one image on top of another while employing subdued halftones. A specific area of this image (Fig. 4.14) will be magnified to ensure the accuracy of the foregoing.

Fig. 4.14. Height difference.

4) Stereo image correction is done in order to resolve this issue. Image rectification (Fig. 4.15) is the process of combining two image planes into one plane so that all of the epipolar lines are perpendicular to the abscissa axis and have the same ordinates in both pictures. Then, using the previously prepared stereo parameters, taking into account bevels, transformations, and the like, undistorted and corrected versions of the input images are returned. Typically, this sub-clause addresses the problem of camera placement errors by making the cameras' focal positions parallel.
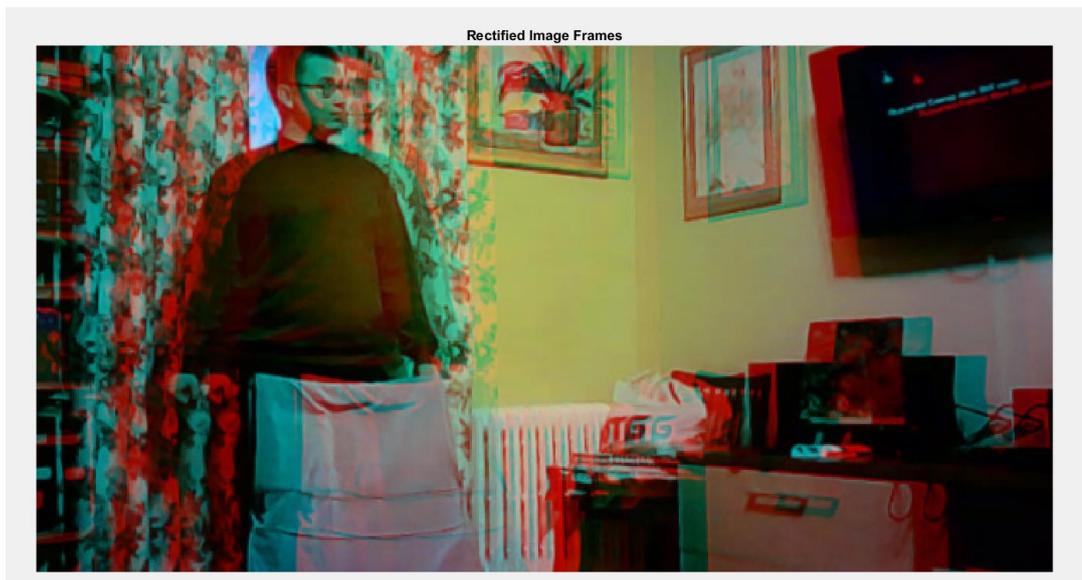
Fig. 4.15. Rectified image frames

It is important to note that because of this feature, the 3D impression is already produced. If the spectator takes special 3D glasses, will see the image in three dimensions, just like being watching a movie in a theater.

5) As can be seen in Fig. 4.15, the conclusion is reached with some image cropping. This is because, when trying to correct images, an appropriate position and angles of rotation of the camera are pursued, but black zones or holes in the plane appear (Fig. 4.16) and have a negative impact on subsequent work; as a result, the image is cropped to the size that is accurate (valid parts).

Fig. 4.16. Uncropped rectified image.

6) The disparity map cannot be used with a color picture; instead, it operates on grayscale, which is the color mode of images that are shown in a table as standards of white brightness. This necessitates the conversion of RGB to Gray (grayscale). Most frequently, a stepped picture of a consistent succession of neutral-gray field optical densities is employed.

7) This step is critical since the Disparity Map only displays a basic colored analog of an object's position at a distance at this point. In other words, regardless of the position, the item is highlighted in a single hue according to the distance, i.e., the closer, the lighter its shade will be (Fig. 4.17).
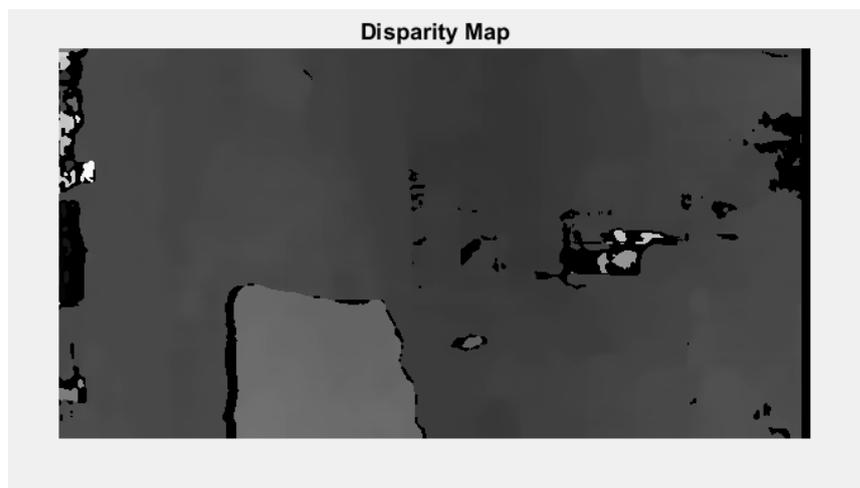
Fig.4.17. Grayscale disparity map

In studies, it was discovered that when transparent entities or those with mirror bases emerge in some regions, sections of them may not be exhibited owing to their structure. This may be caused by the uniformity of the color of the item itself or by the glare of light on the camera. A range of parameters with the dimension [0 64] was added to the scenario to make it more aesthetically pleasing and easy to analyze (Fig. 4.18). The number of colors can be manually adjusted or left blank; in any case, the lowest and maximum amounts are allowed based on the kind of recorded format; for uint16, for example, this parameter is [0 65535].
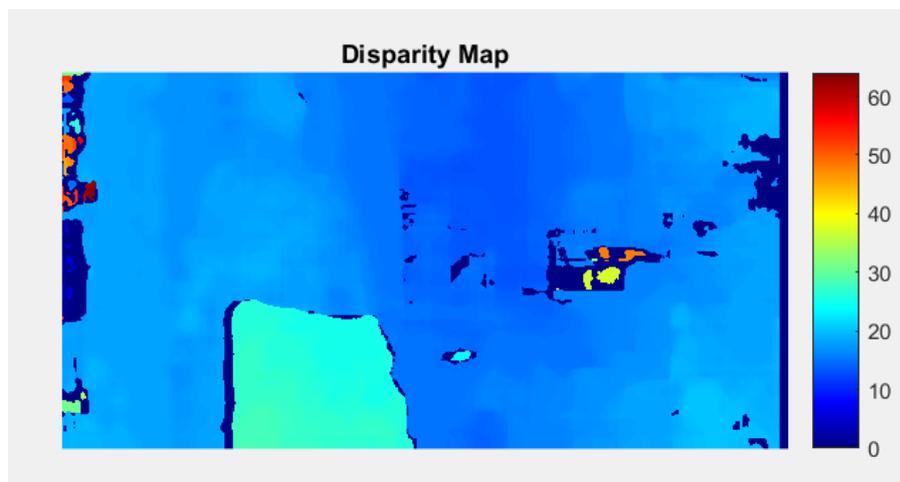
Fig. 4.18. Colored disparity map

8) Scene reconstruction is done by attaching stereo parameters to the disparity map; in other words, if a typical disparity map model is received, it is used to produce a scene since it already has all of the world coordinates for the points that are saved in the picture. The parameter input in this situation must be the same as that used for correction, accounting for any stereo image adjustments that were made prior to the matching disparity maps. The data must then be recast in meters because the initial data was stored in millimeters after receiving the data with each point's depth; to do this, all internally values are scaled by 1000. The 3D scene can then be shown on the screen, but since only data from the Disparity Map arrangement is sent there and, as previously mentioned, a picture is only shown there in halftones, it won't be very clear to a person. To make the project more

realistic, values from one of the color pictures are added to the code. Additionally, it would be beneficial to include trash collection when presenting on the display, as it was only interested in scenes that were no larger than 2 meters in width, height, and length, with a maximum of 5 meters. This was because the parameter was configurable and used according to individual circumstances.

9) Volumetric representations are filled simply to accomplish the display task, and this allows for numerous operations like increasing rotation and grabbing a point. The cameras ought to be perpendicular and not experience focus shifts during recording. Examples of many rooms are Figs. 4.19, 4.20, 4.21, and 4.22.
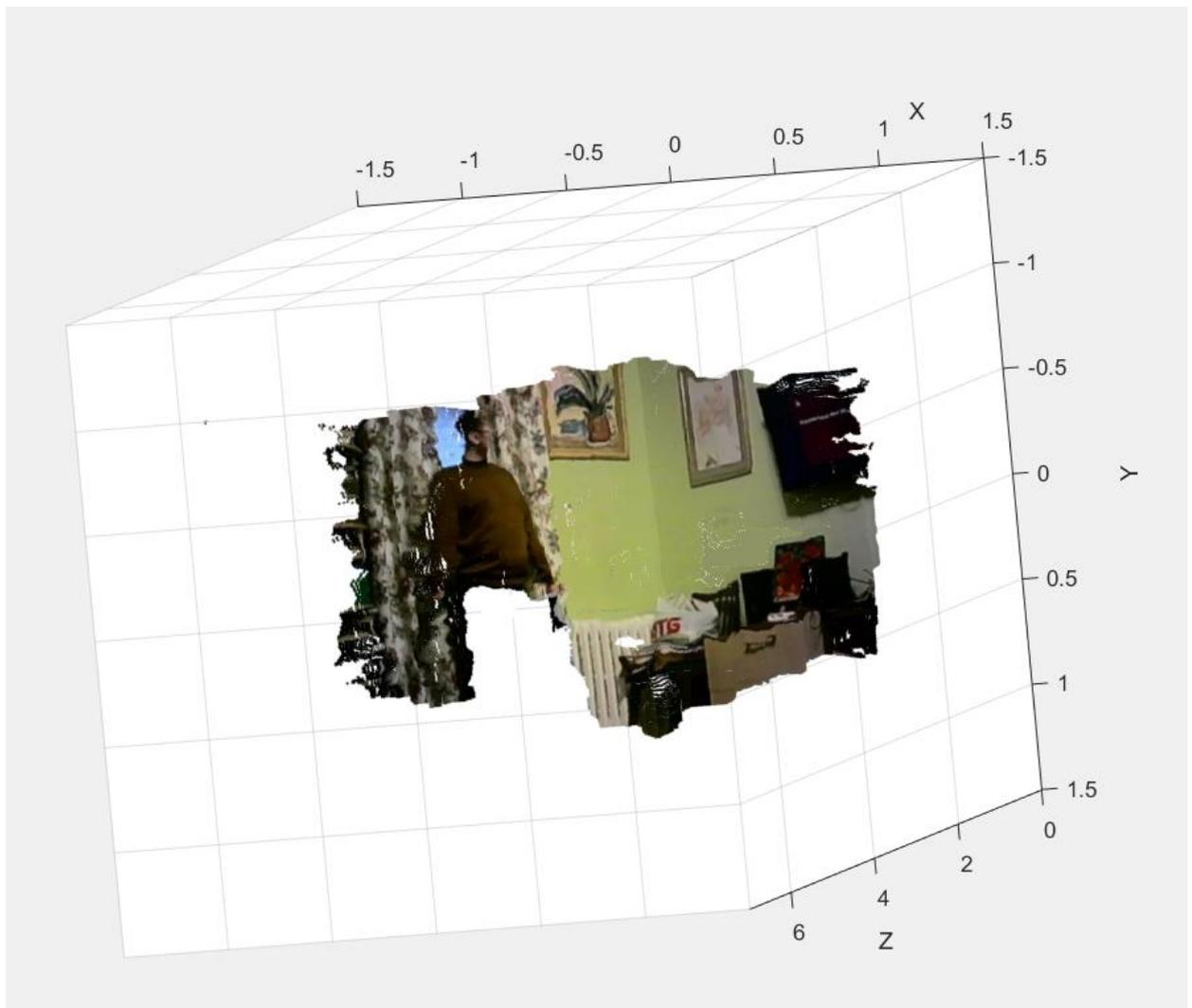
Fig. 4.19. Example 1
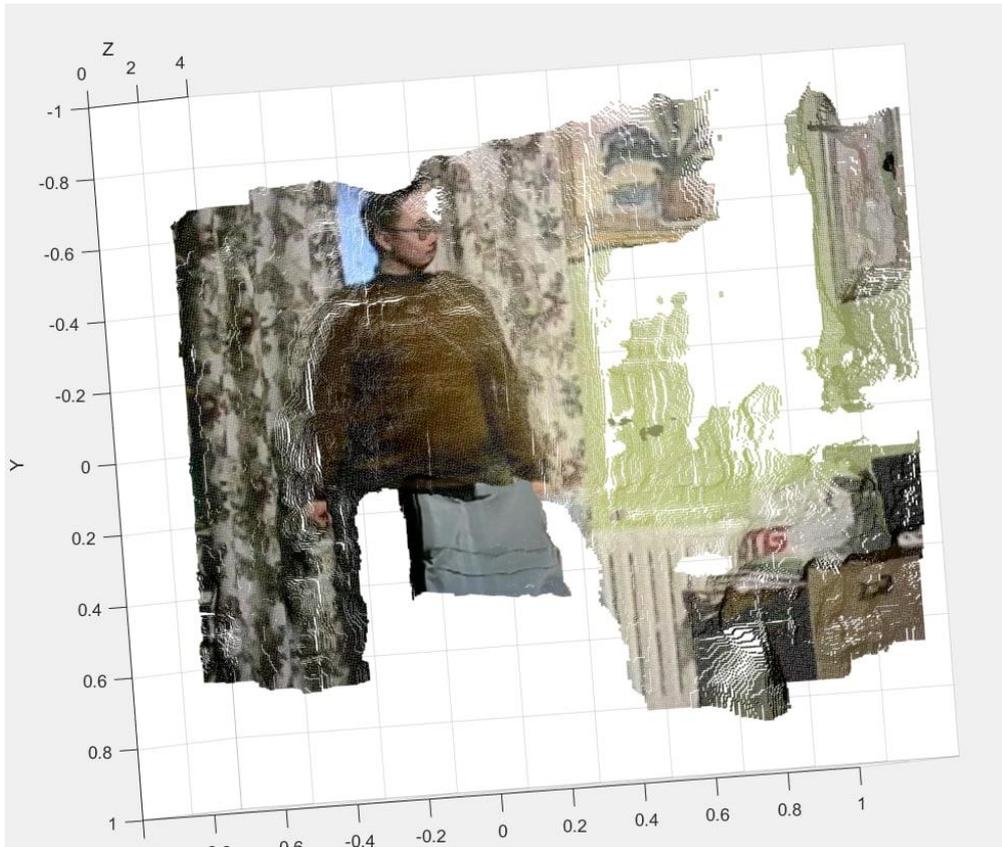
Fig. 4.20. Example 2
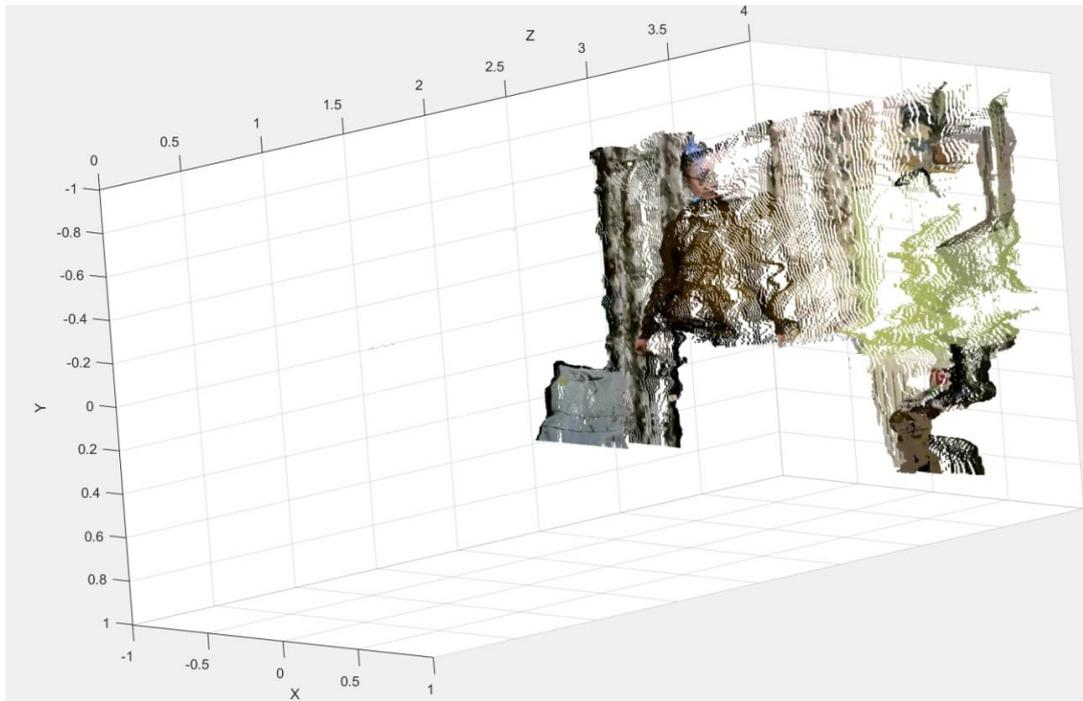
Fig. 4.21. Example 3
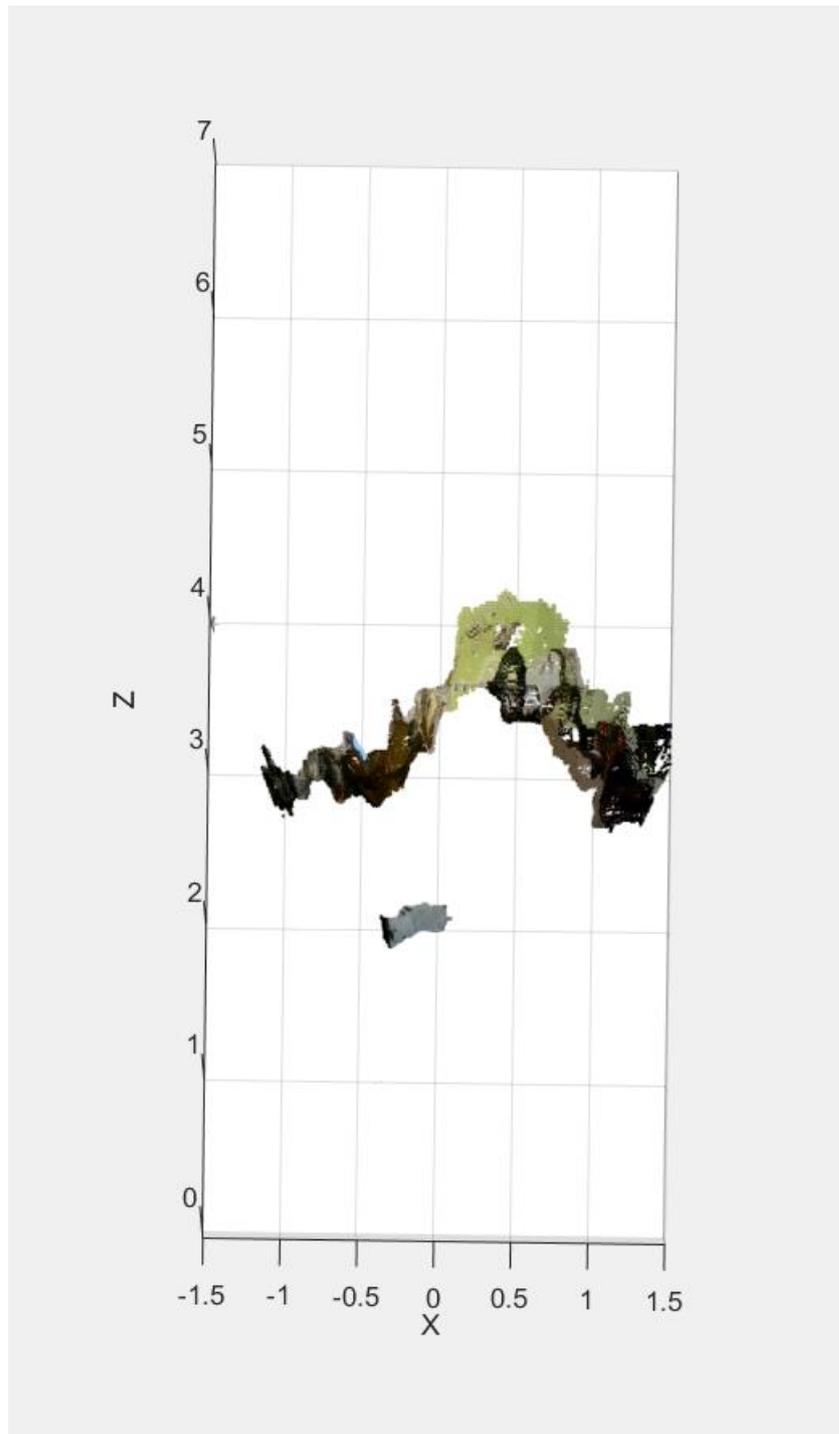
Fig. 4.22. Example 4

When the scene rotates, it is possible to see a three-dimensional image with cutouts in the areas where the bodies are obscured by one another or are not in the field of view, making it easy to determine the approximate distance to objects from these situations.

The implementation of streaming is similar to the algorithm for images described above, with the exception that it operates on the idea of scripting images from the right and left cameras through a lower limit of time delay and produces only a 3D scene because this process is looped and is called many times when calling any other, making it difficult to view and comprehend. Rotating the scene is also unacceptable because the figure is built many times and re-displays (Fig. 4.23).
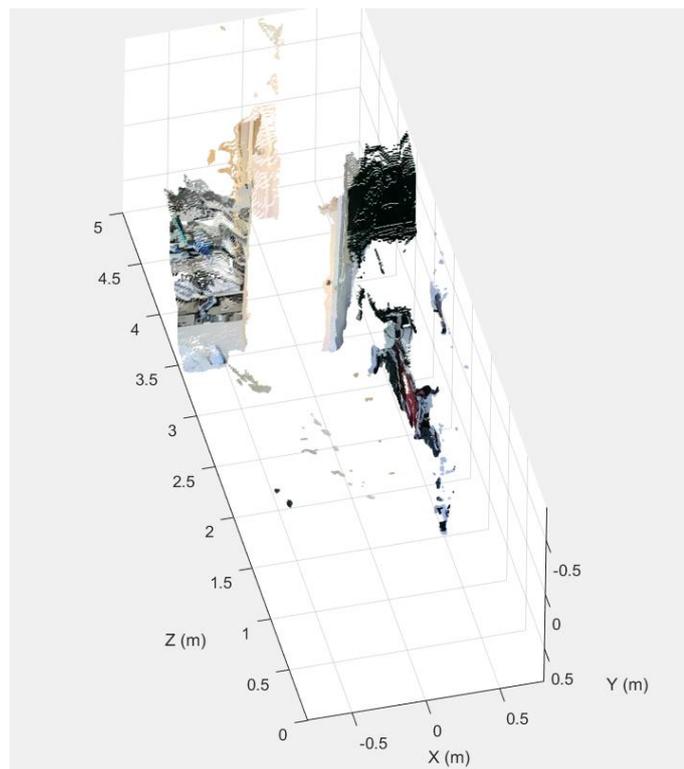
Fig. 4.23. 3D scene streaming

The cycle's operation can be terminated if required, and the data that was captured at the final minute is preserved for use with it in the future.

4.6. Filtering image pairs to improve the display of a 3D scene

Since during shooting there are a number of disadvantages related to lighting, reflection of light from surfaces, which directly worsens the image display result, there is a need to create internal filtering of images before sending them for processing.

For this, the built-in functions of image filtering in Matlab will be used and the best type of filtering that can give the best result will be selected.

Currently, the Matlab program offers the following types of filtering of 3D images: Gaussian filter of 3D images, 3D median filter and 3D box filter.

4.6.1. Gaussian image filter

The Gaussian blur filter is quite often used by itself or as part of other image processing algorithms. Next, a method will be described that allows to obtain blur with a speed that does not depend on the blur radius, using filters with an infinite impulse response.

So, let the original image be set to brightness $x(m, n)$. Gaussian blur with radius r is calculated by the formula:

$$y(m, n) = \frac{1}{2 \cdot \pi \cdot r^2} \sum_{u,v} e \cdot \frac{-(u^2 + v^2)}{2 \cdot r^2} \cdot x \cdot (m + u, n + v) \qquad (4.1)$$

The summation limits of u and v can be chosen plus or minus a few sigma, i.e. radius r, which gives the complexity of the algorithm of the order of $O(r^2)$ operations per pixel.

In order to evaluate the result of the filter, an old image that described a 3D scene will be used (Fig. 4.19). Thus, after processing the image with a sigma value equal to 2, the filtered image takes on the following appearance (Fig. 4.24).

As can be seen from the ratio, the reflection loses sharpness, but the objects become more smoothed, which gives a number of advantages in terms of the fact that reflections from surfaces and glare from light are lost to a greater extent.
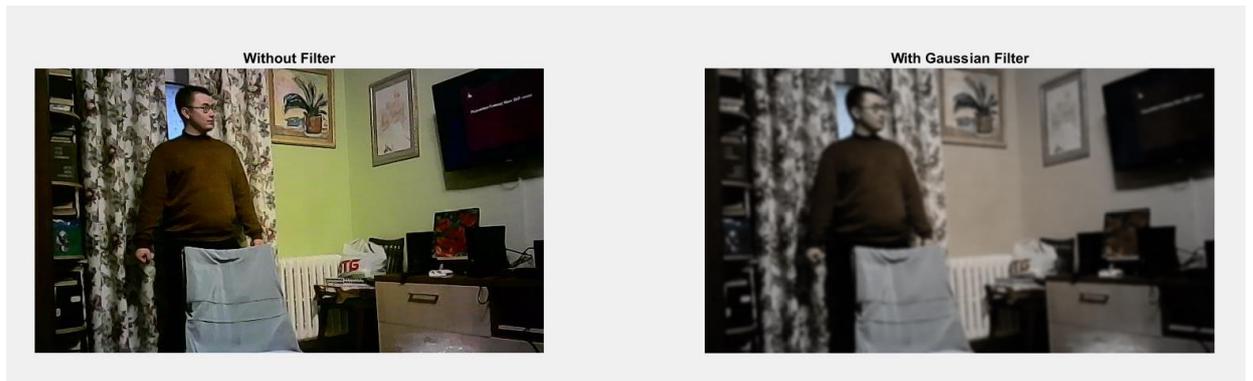
Fig. 4.24. Images without and using a Gaussian filter

In order to understand that the filter improved the result of forming a 3D scene, the program was launched using it (Fig. 4.25).
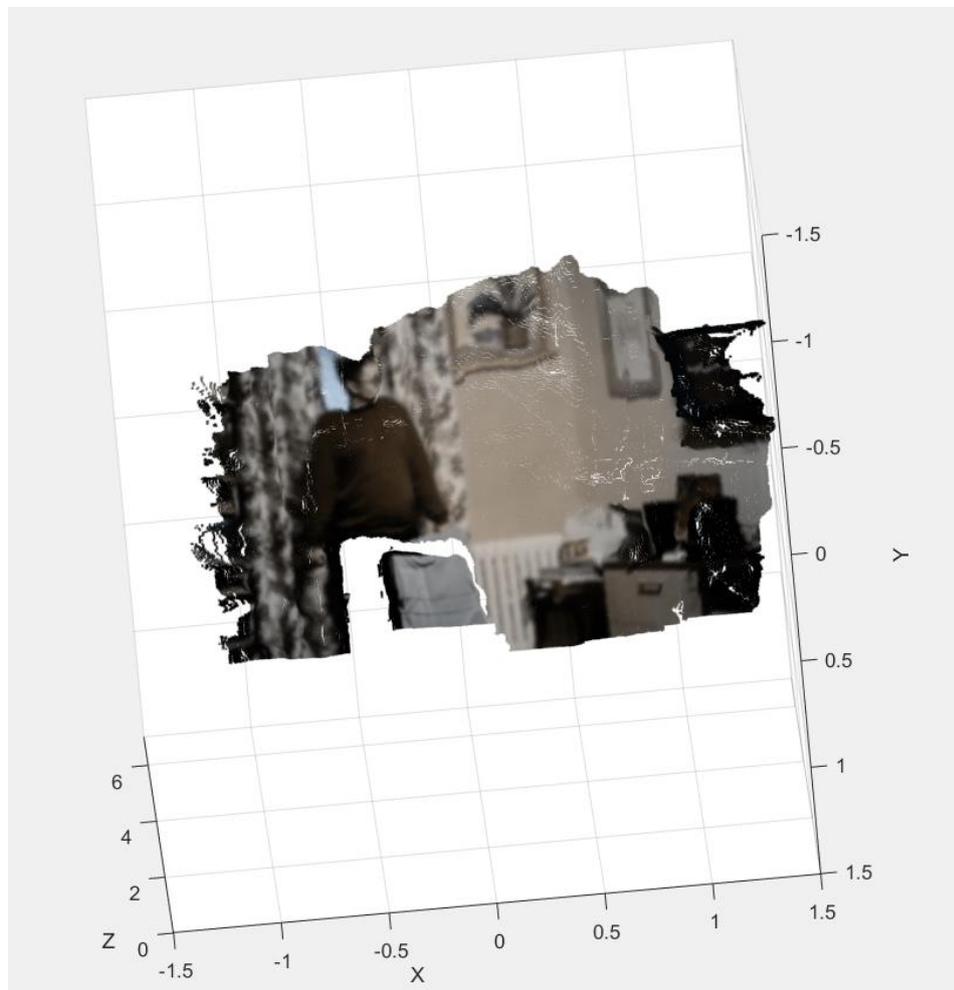
Fig. 4.25. 3D scene using the Gaussian filter

As it can be seen from the image, the filtering helped to eliminate the loss of some components due to light reflection and is reasonable to use as the result is improved.

4.6.2. Median image filter

The median filter is the most significant interference suppression encountered in the development of this project. It is just right for eliminating all kinds of small scattered inclusions, but it is not an ideal tool because passes slightly larger areas of concentrated interference.

The median filter is a 3×3 pixel sliding window. At the input, it acquires 9 values (pixels), and at the output it gives one. The median filter works as follows: it sorts the input data (pixels) in ascending order and outputs the middle result (median).
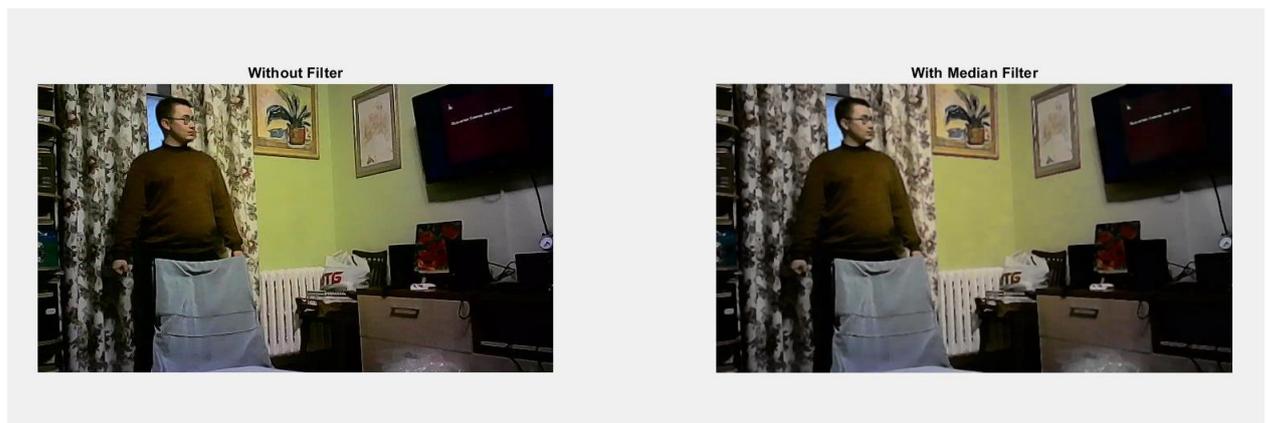
Fig. 4.26. Images without and with the use of a median filter

As it can be seen, there have been no fundamental changes, because of this the result of constructing a 3D scene does not improve (Fig. 4.27), as noise will remain in the image.

It should be noted that after using this method, clearer boundaries of certain objects appeared.
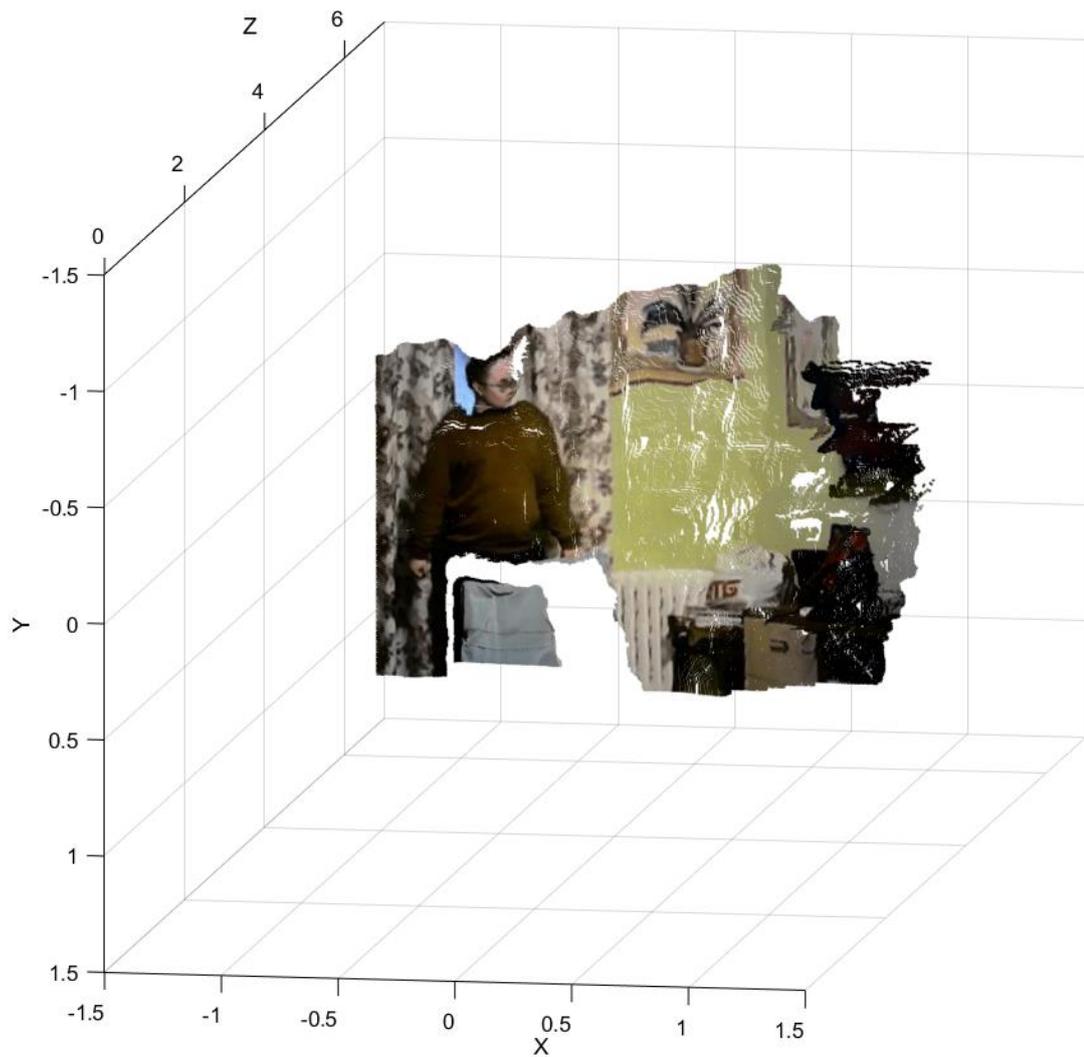
Fig.4.27. 3D scene using a median filter

4.6.3. Box filter

Imboxfilt accomplishes filtering either use the integral image filtering or convolution-based filtering, choosing which filtering method to employ based on an inherent algorithm.

Examples of mapping are shown below in Fig. 4.28 and Fig. 4.29.

This filter processes the image better than the median filter, but worse than the Gaussian filter, it can be used, but with the addition of additional internal settings.
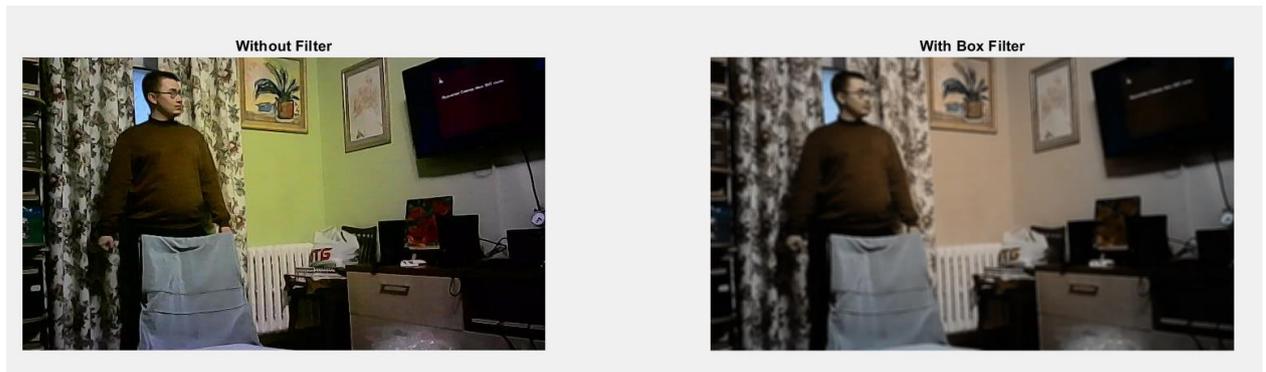
Fig. 4.28. Images without and using a box filter
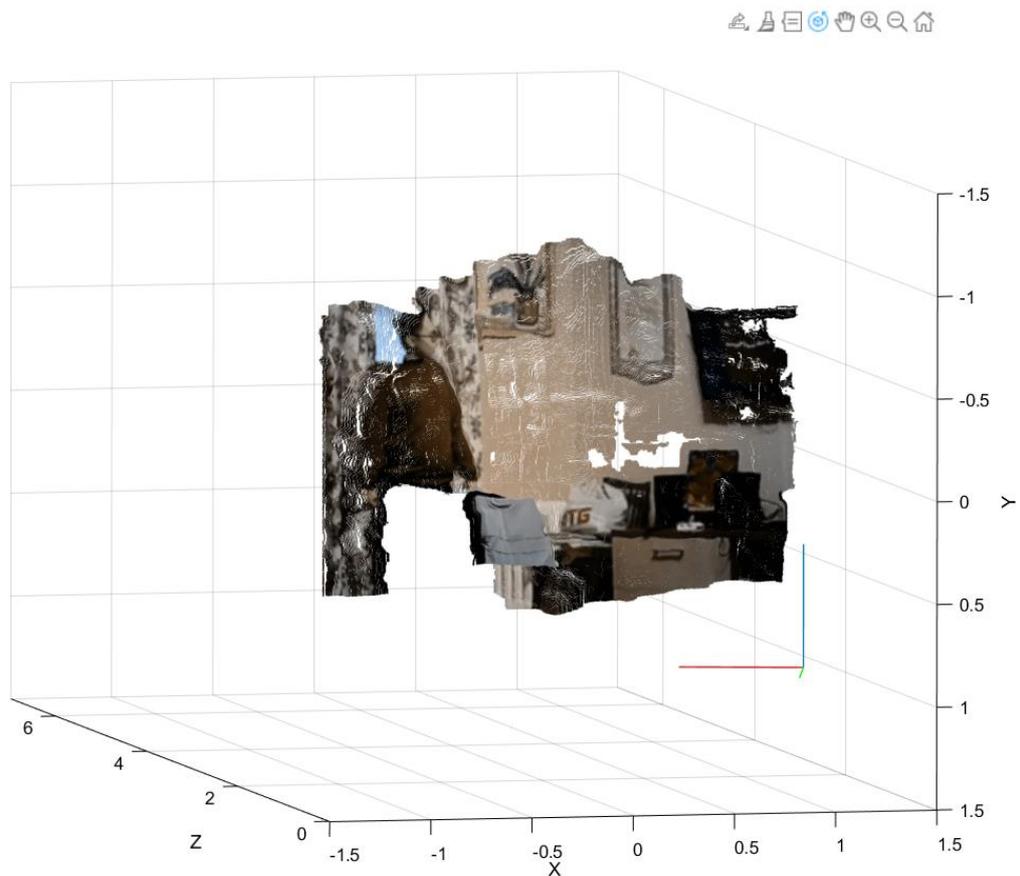
Fig.4.29. 3D scene using the box filter

4.6.4. Combined filtering method

Since during experimental research it was found that the best accuracy results are characterized by the Gaussian and median filters, it was decided to combine them into one (Fig. 4.30, Fig. 4.31) and find out whether this affects the final result.
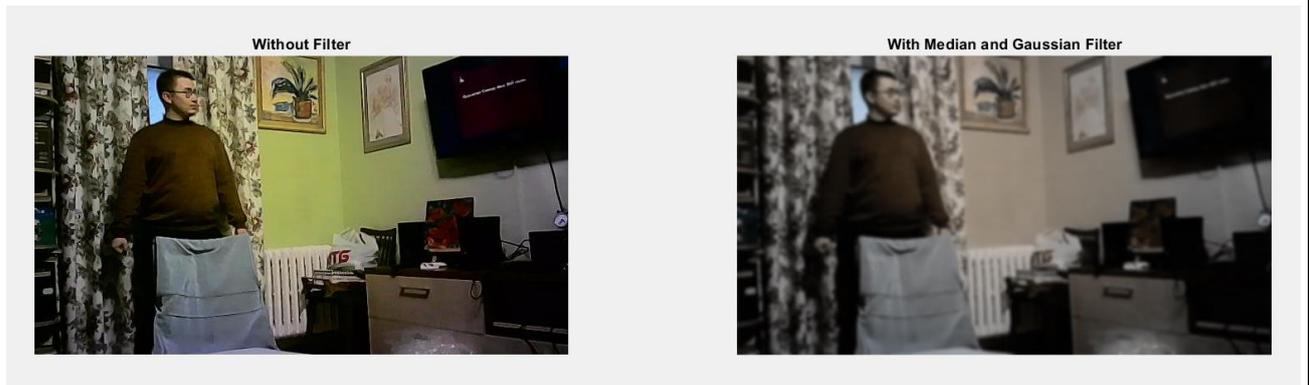
Fig.4.30. Images without and using a combined filter
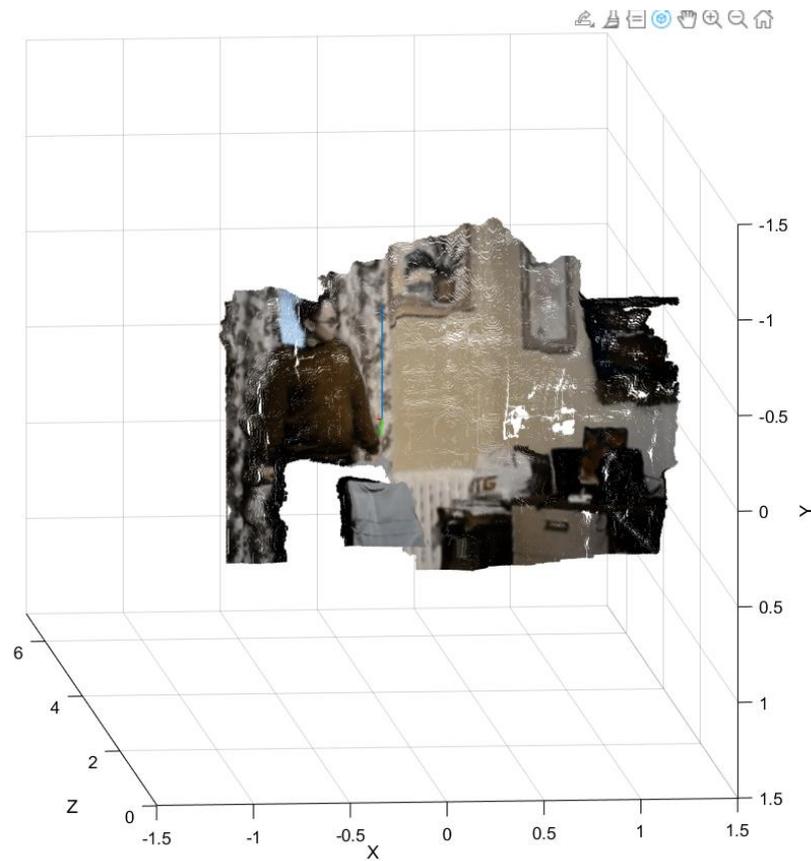
Fig.4.31. 3D scene using a combined filter

The study confirmed that combining these methods with a sigma equal to one in the Gaussian filter gives the best result and will be used in the future.

4.7. Using the model on the external environment

Now it will be considered what results the system gives when observing objects on the street. In the course of the research, a number of shortcomings were

revealed, which are difficult to get rid of. Since most of the bodies located outside do not have a stable or reflective surface, the incorrectness of the initial data was found. For example, if it is necessary to display objects that are at a great distance, it is necessary to change the design model, since the distance between the camera lenses is directly proportional to the distance of the observed object, if it is necessary to find out the conditional distance to a body that is more than 20 meters away, it is necessary to anew recalibrate cameras with a pre-increased distance between optical lenses, the same algorithm occurs in the opposite case.

The first object of observation was a car that was located at a distance of 3.20 meters. For understanding, the initial image and the image after filtering are shown (Fig. 4.32).
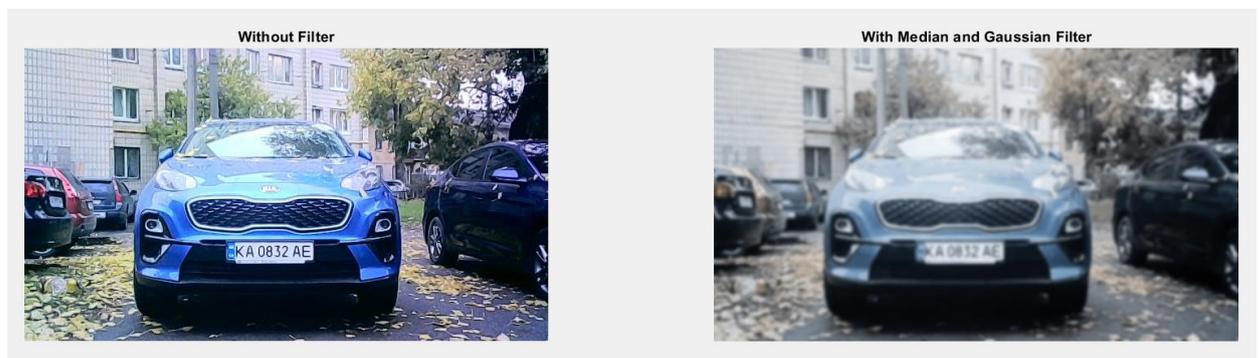
Fig. 4.32. Initial and filtered view of the image for further processing

Below is a 3D scene of the car and the surrounding objects, it should be noted that the covering of the car created significant obstacles for accurate reflection and as you can see in the image, some of the surfaces were completely cut off, because the mirror principle was formed, and the focal lens could not determine the point of the object. It can also be seen that the shooting took place in a triangular shape, because of this, some of the objects could not be reflected, because the focal distance was not defined for them, for example, the rear part of the right neighboring car was well reflected, and the front part was not. From the received data, it can be noted that the cameras worked well at a distance of up to 5 meters, then the image began to cut off.
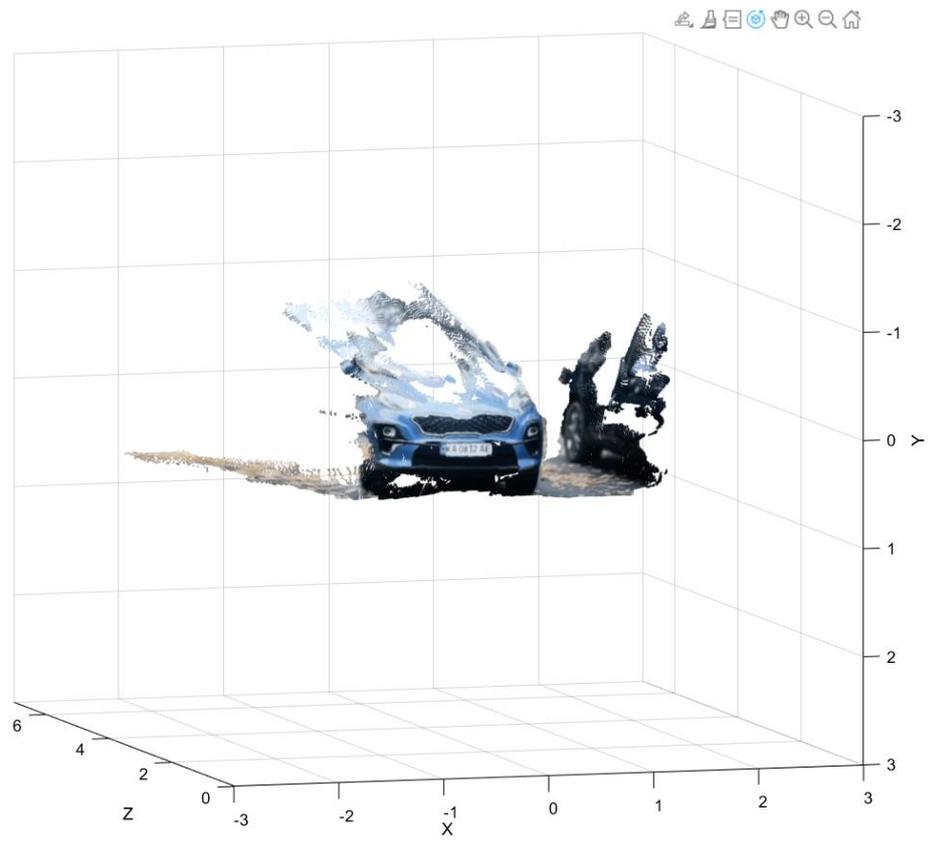
Fig. 4.33. 3D scene of the environment, example 1
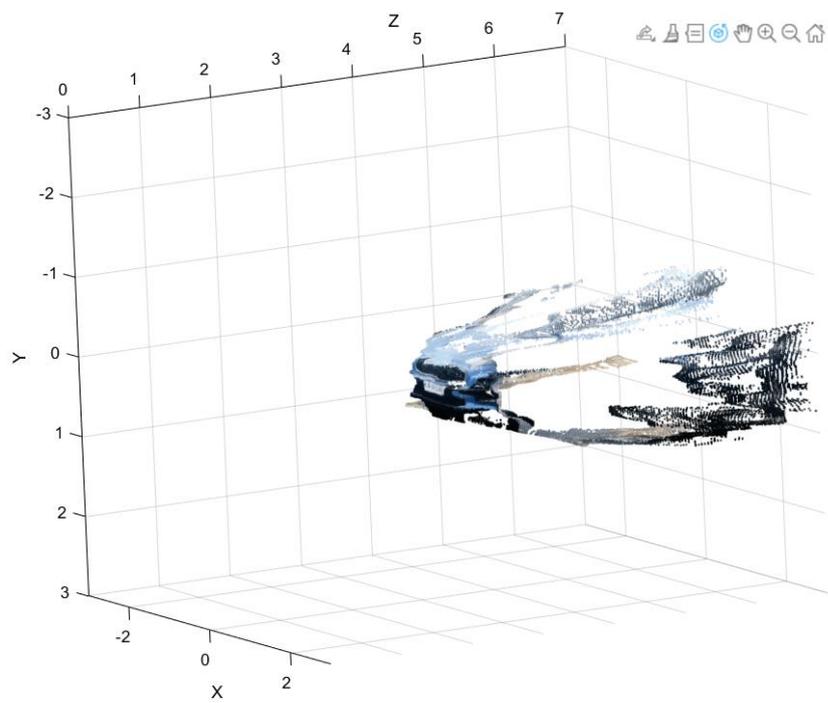
4.34. Example 2

A few more examples of displaying cars will be given below.
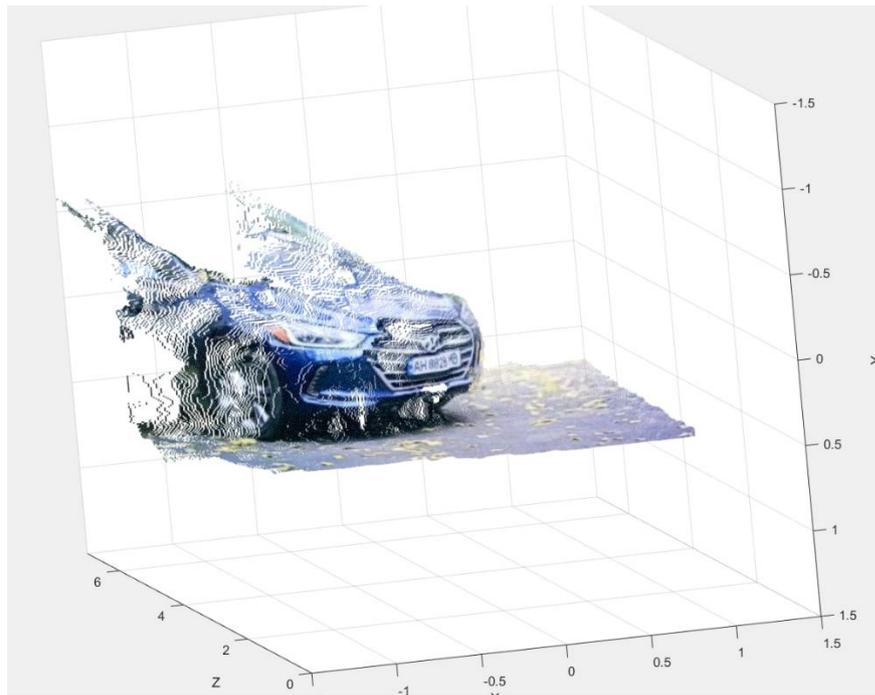
Fig. 4.35. Example 3
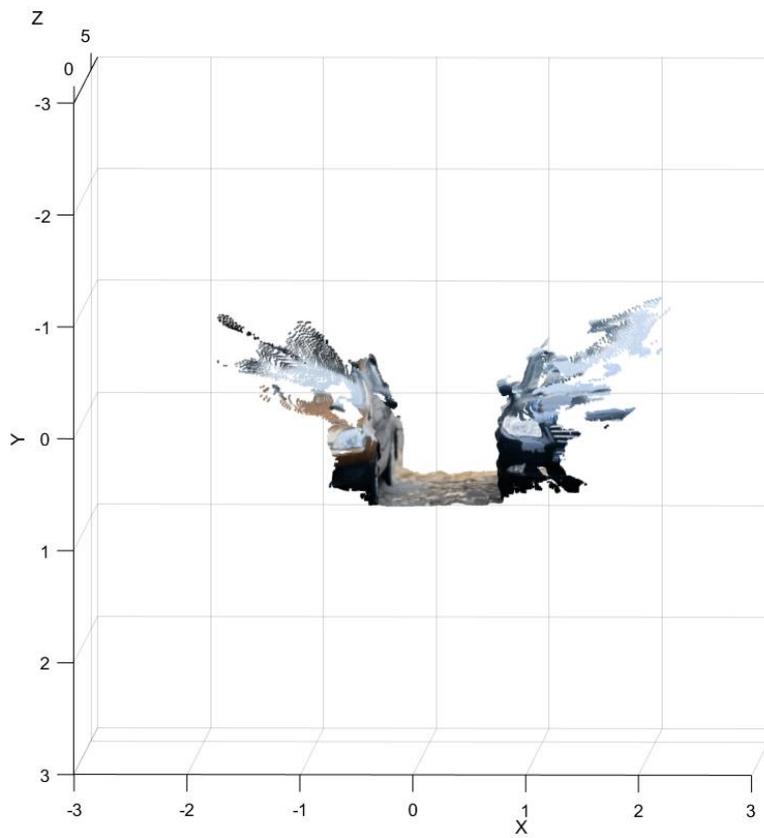
Fig. 4.36. Example 4

Attempts have also been made to depict pedestrians in the courtyard, examples of which are given below.
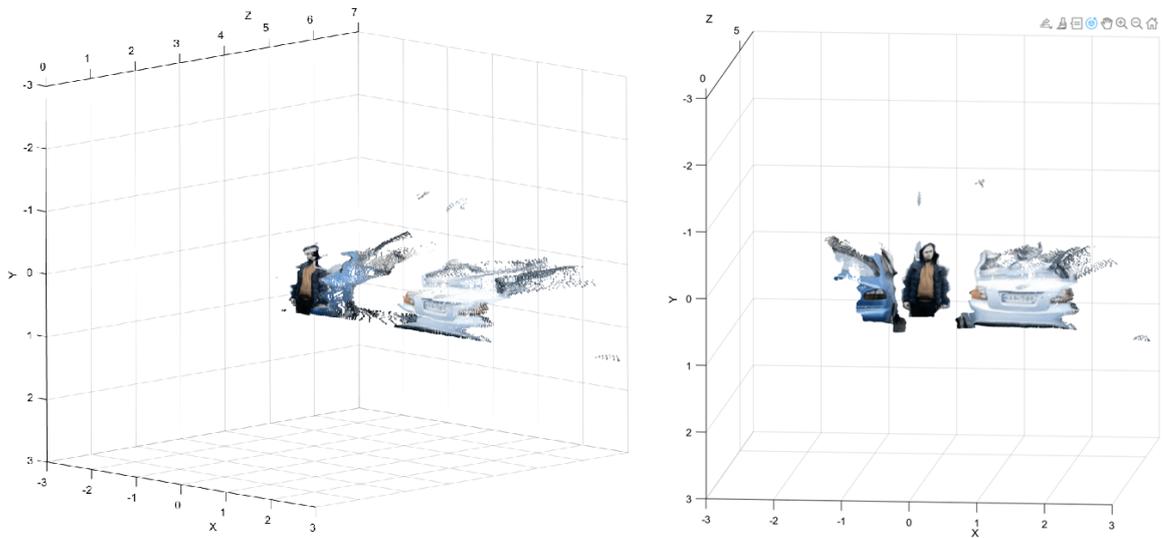
Fig. 4.37. Display of pedestrians in space, example 1
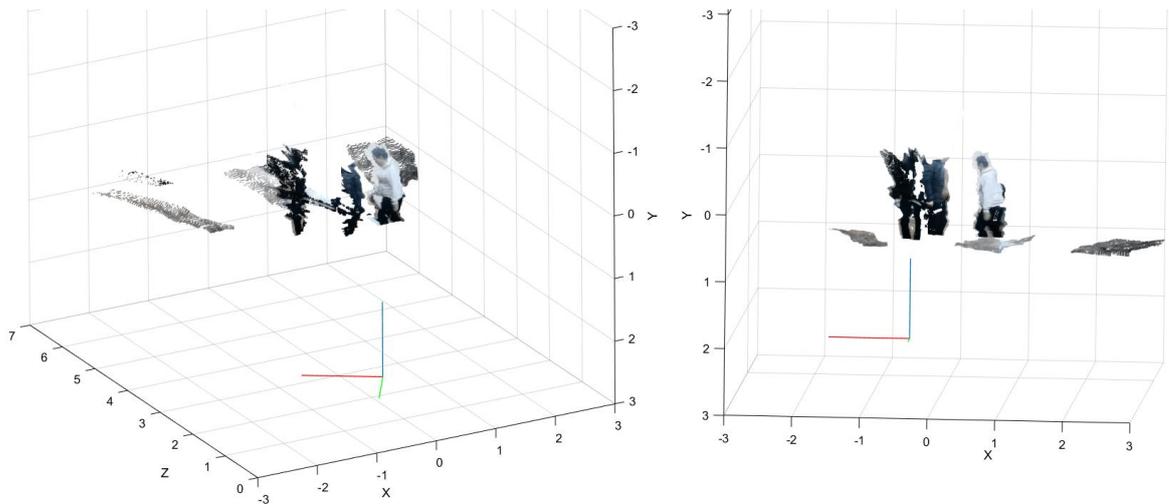
Fig. 4.38. Example 2

In the course of experiments, it was found that the optimal distance for reflection without distortion is up to 5 meters, since the initial setting was made for shooting indoor spaces, and not for outdoor environments, but this is easily corrected by increasing the distance between the lenses and re-calibrating.

Unfortunately, this model is not quite suitable for use on the street, but gives quite true data at the output.
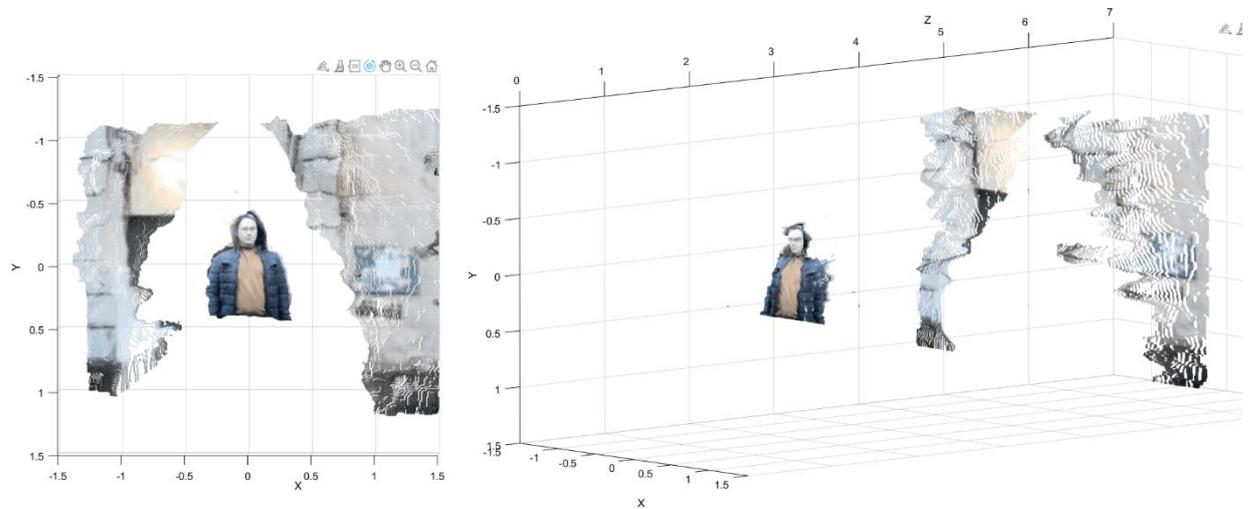
Fig. 4.39. Example 3

4.8. Analysis of the accuracy and operability of the working structure

An examination was conducted taking into consideration the precise calculation of the range to objects to ensure the installation's satisfactory functioning. The body was initially positioned at a varied distance of up to 20 centimeters and moved up to a distance of 4 meters. A distance of more than 5 meters is considered only in cases of outdoor use. Table 4.2 shows the results of indoor measurement accuracy.

It is demonstrated that after presenting both actual data and program data, it was feasible to determine the absolute and relative inaccuracy of the task. The ratio of values is depicted on the graph in Fig. 4.40.

Table 4.2. Accuracy check of depth map in real and observed range

Undisturbed environment			
Real distance, m	Distance measured by the depth map, m	Absolute Error, m	Relative Error, %
1	1,08	0,08	8,00
1,2	1,28	0,08	6,67
1,4	1,32	-0,08	5,71
1,6	1,51	-0,09	5,63
1,8	1,87	0,07	3,89
2	2,048	0,048	2,40
2,2	2,28	0,08	3,64
2,4	2,47	0,07	2,92
2,6	2,59	-0,01	0,38
2,8	2,88	0,08	2,86
3	3,1	0,1	3,33
3,2	3,15	-0,05	1,56
3,4	3,42	0,02	0,59
3,6	3,66	0,06	1,67
3,8	3,77	-0,03	0,79
4	4,08	0,08	2,00
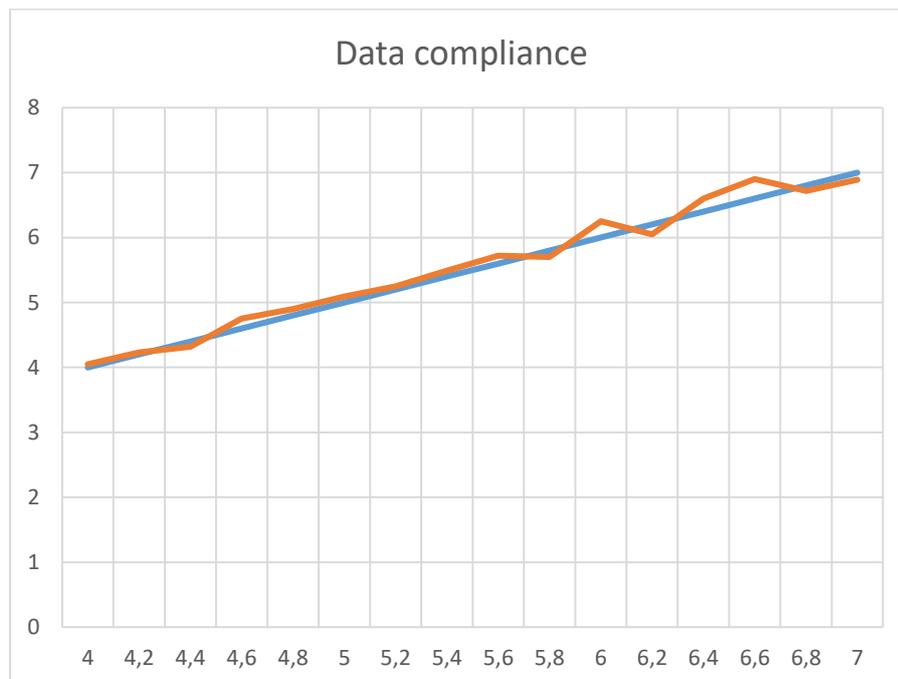
Fig. 4.40. Data comparison

For devices intended to measure the distance to an object, the model can be considered sufficient as long as the error does not exceed 5%.

Additionally, the accuracy findings may differ based on the location of the device because the environment might have an impact for a variety of reasons, including pressure, surface reflection, light, etc.

Table 4.3. Accuracy check of depth map in real and observed range

Disturbed environment			
Real distance, m	Distance measured by the depth map, m	Absolute Error, m	Relative Error, %
4	4,05	0,05	1,25
4,2	4,23	0,03	0,71
4,4	4,32	-0,08	1,82
4,6	4,75	0,15	3,26
4,8	4,9	0,1	2,08
5	5,09	0,09	1,80
5,2	5,25	0,05	0,96
5,4	5,49	0,09	1,67
5,6	5,72	0,12	2,14
5,8	5,7	-0,1	1,72
6	6,25	0,25	4,17
6,2	6,05	-0,15	2,42
6,4	6,6	0,2	3,12
6,6	6,9	0,3	4,55
6,8	6,72	-0,08	1,18
7	6,89	-0,11	1,57

As in this work, the system was also used externally, an analysis of the accuracy of data was carried out externally (Table 4.3), the object shown in Fig. 4.32 was selected for measurements.

The dependence graph of measurement errors is shown in Fig. 4.41, it shows that as the distance to the point increases, the error increases, which is not acceptable for developers, this is solved thanks to the calibration methods described earlier.
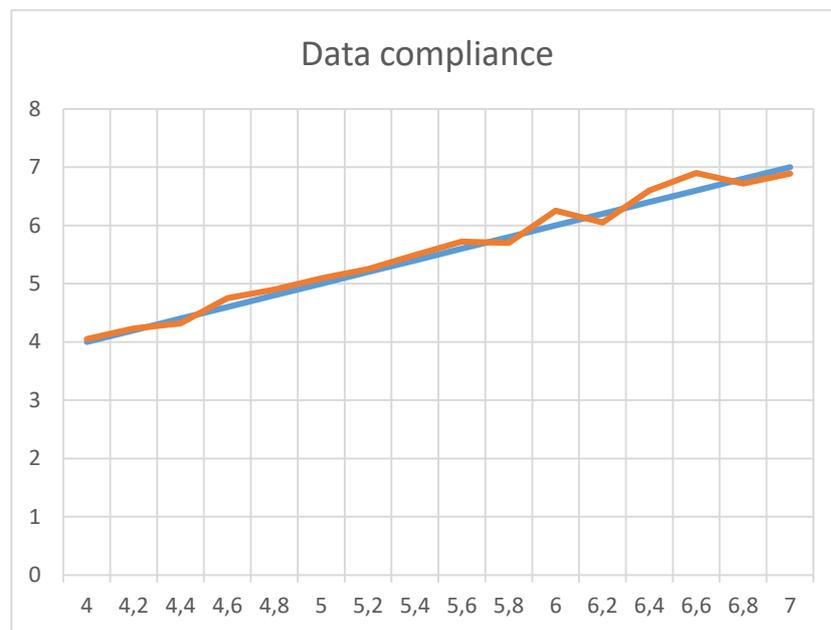
Fig. 4.41. Data comparison

4.9. Power consumption being used during computing

Nowadays, a program's speed plays a significant role, therefore a developer must be aware of the prices and hardware requirements for reliable, high-quality work. Since the application is primarily designed for live filming, which may place a heavy burden on the CPU, video card, and needed RAM, it is crucial that it function without errors and stuttering.

The block diagram of the streaming 3D scene's output device has the appearance displayed in Fig. 4.42.
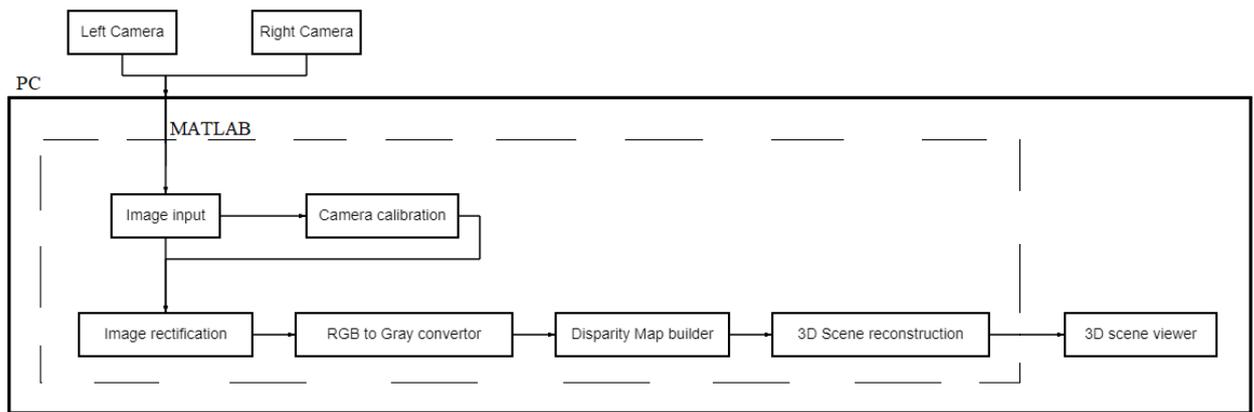
Fig. 4.42. Block diagram of a streaming 3D scene's output device

To begin, the features of PC (Table 4.4) and the cost of the utilised RAM will be discussed in order to determine how quickly the software can operate. Even if the system requirements for the PC were not the finest, it was sufficient.

Table 4.4. Characteristics of PC

PC name	Lenovo Ideapad 320 15i-KB
CPU	Intel Core i3-7130U CPU 2,7 GHz
Number of Cores	2
Logical processors	4
Installed RAM	12 Гб
Video adapters	Intel(R) HD Graphics 620 NVIDIA GeForce 940MX

Then, an experiment was conducted to determine the system load when processing a single image. It was found that the CPU load during this experiment was not greater than 20%, and the time taken did not exceed 3 seconds. This number may seem excessive, but keep in mind that there is a search and distribution of the image from the PC storage device to the program, then it is already transformed with multiple viewing methods, for a more outstanding example,

before it is converted. From perspective, these outcomes are fairly respectable for just working with a single image. By the way, the memory expenses were just 900 MB, which is really low for today's demanding apps, which is undoubtedly a significant bonus.

Let's go right to the video streaming task now. The memory costs will undoubtedly be considerably higher than when dealing with a single image, but there is a huge benefit in that the loop will run much more quickly than the program code (Appendix C) does for a single scene. This occurs because calling and displaying extraneous information are unneeded, resulting in the same screen is eventually just overwritten.

The tic toc function (Fig. 4.43) in MATLAB, which operates on the clock principle, toc – end, tic - start of timer, was used to determine the precise time of passing one cycle.

```
Elapsed time is 19.528022 seconds.
Elapsed time is 19.943868 seconds.
Elapsed time is 20.364600 seconds.
Elapsed time is 20.785838 seconds.
Elapsed time is 21.225324 seconds.
Elapsed time is 21.696200 seconds.
Elapsed time is 22.144253 seconds.
Elapsed time is 22.715397 seconds.
```

Fig. 4.43. A tic-toc function that records a stream program's cycle time

After the experiment, it was established that the scene is rewritten approximately twice per second. This is an excellent result considering that the camera itself does not shoot at a rate of more than 4 frames per second, and that when processing was taken into account, the rate was only two frames per second.

Moving on to the needed memory, it was surprising to see that the spent memory stayed at the same level even though the CPU was starting to load 10% more than it normally would. This is a decent outcome, but can't say whether

working on other PCs won't be more stressful because the system hasn't been tested on those particular machines.

CHAPTER 5. LABOR PROTECTION

5.1. Introduction

Video surveillance using stereo vision is a system of complete control and protection of your organization from all kinds of unpleasant situations or determination of distances, volumes or sizes to objects in production. Commercial objects need a whole video surveillance system. The video surveillance system in labor protection is not only cameras, but a whole software and hardware complex. After all, in addition to shooting video, it performs the task of processing and transmitting information to the control device, which is a personal computer that performs the function of saving and transfer of information to other devices.

Areas of use can be spread across many industries:

- these are various hard-to-reach objects - installation work at height or depth;

- it can be simply epidemiological conditions - most of the information systems work remotely;

- these are ordinary workplaces, during the daily work of the eyes, the performer's vision becomes more beautiful than ever. And movements, at first glance, practiced to the point of automatism can fail. And it is not far from an industrial injury.

Thus, in all three cases, the video surveillance system has the most effective effect on the safety of the labor process.

Video surveillance systems are of several types [24] depending on the object of classification:

- Digital or analog systems;

- Wired or wireless;

- Network or autonomous;

ACIC DEPARTMENT		NAU 22 1094 000 EN			
			N.	Page	Pages
Performed	O. S. Sych	IMAGE DEPTH DETECTION SYSTEM		111	160
Supervisor	M.P. Vasylenko				
Consultant	O.O. Kozlitin		225 151		
S. controller	M.K. Filyashkin				
Dep. head	V.M.Sineglazov				
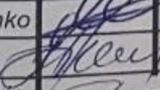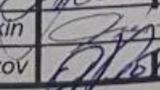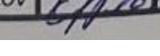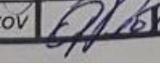
- Domestic, hidden, street or security.

The system may have the following elements:

- The video surveillance cameras themselves: transmission and fixation of the image;

- Rotating devices for cameras: expand the viewing angle, allow you to move the camera up, down, right, left. Thanks to these devices, the viewing angle is significantly increased;

- Camera lenses: increase the range of work, help to adapt to the specific environment of the selected object;

- Devices for processing incoming signals: signals are transmitted to monitors;

- Recording devices: record a continuous streaming image;

- Video monitors: you can purchase color or black-and-white monitors;

- Switches (matrix): connect video cameras to a video recorder, multiplexer or monitors;

- Devices for printing images: not necessarily devices for buildings that install video cameras as a preventive measure, but where unpleasant situations happen quite often - an irreplaceable thing.

Depending on the type of object, the composition may change.

5.2. Analysis of working conditions at the workplace

Since the worker is required to prepare the installation itself before starting work, i.e. advance calibration and adjustment of the software device and the workplace for comfortable use, there are a number of requirements for preparing the workplace [25] before starting work.

1) Inspect the workplace and put it in order; make sure that there are no foreign objects on it, all equipment and PC units are connected to the system unit with connecting cords;

2) Check the reliability of installing the equipment on the desktop. The monitor should not stand on the edge of the table. Turn the monitor so that it is

convenient to look at the screen - at a right angle (not from the side) and slightly from top to bottom; at the same time, the screen should be slightly tilted - the bottom edge is closer to the user;

3) Check the general condition of the equipment, the serviceability of electrical wiring, connecting cords, plugs, sockets, grounding of the protective screen;

4) Insert the plug into the socket and make sure that it holds firmly. It is forbidden to insert and remove the plug with wet hands.

5) Adjust and fix the height of the chair and the user-friendly backrest angle.

6) If necessary, connect the necessary equipment to the computer (printer, scanner, etc.). All cables connecting the system unit to other devices should be turned on and off only when the computer is turned off.

7) Adjust the brightness and contrast of the monitor.

8) Inform the supervisor of all detected malfunctions and do not start work until they are eliminated.

9) An employee who has the appropriate working qualifications, who has passed a medical examination and has no health contraindications, who has undergone labor protection briefings, training and a knowledge check of labor protection requirements, is allowed to work on the use of video surveillance systems.

10) To perform work using technical means of information processing and transmission, it is necessary to study the instructions for operating the equipment, undergo special training and obtain an electrical safety group.

11) An employee who uses video surveillance systems in his work, regardless of qualifications and work experience, must undergo repeated instruction on labor protection at least once every six months.

12) An employee who has not undergone briefings, training and knowledge testing on labor protection is not allowed to work independently.

13) An employee allowed to work independently must know the rules of operation of video surveillance systems. Rules of operation of the applied technical means of information processing and transmission. Rules, regulations and instructions on labor protection and fire safety. Rules for using primary fire extinguishing means. Methods of providing first aid in case of accidents. Rules of the organization's internal labor procedure.

14) An employee who has shown unsatisfactory knowledge and skills in the safe performance of work using video surveillance systems is not allowed to work independently.

15) An employee who is directed to participate in work uncharacteristic of his profession (position) must undergo targeted instruction on the safe performance of future work.

16) When working with a video surveillance system, an employee may be adversely affected, mainly by the following dangerous and harmful production factors:

- Prolonged static tension of the muscles of the back, neck, arms and legs, which can lead to static loads;

- Overvoltage of auditory and visual analyzers when receiving and transmitting information;

- Electric current, the path of which in the case of a short circuit can pass through the human body.

17) To prevent the possibility of a fire, the employee must comply with fire safety requirements himself and prevent violations of these requirements by other employees.

18) The employee is obliged to observe labor and industrial discipline, the rules of internal labor regulations.

19) In case of illness, poor health, or lack of rest, the employee should inform his immediate supervisor about his condition and seek medical help.

20) If an accident happened to one of the employees, the victim must be given first aid, report what happened to the supervisor and preserve the situation of the event, if it does not create danger for others.

21) The employee, if necessary, must be able to provide first aid, use a first aid kit.

22) An employee who violates or fails to comply with the requirements of the occupational health and safety instructions is considered a violator of industrial discipline and may be subject to disciplinary liability and, depending on the consequences, to criminal liability; if the violation is related to causing material damage, then the culprit may be held materially liable in accordance with the established procedure.

5.3. Development of labor protection measures

Users may be exposed to the following types of hazards while working with computer equipment:

- electric shock;

- energy hazard (caused by short circuit: burns, electric arc);

- danger of fire;

- thermal hazard (effect of high temperatures due to heating of structural elements);

- mechanical danger (injuries due to falling, action of moving parts, cut by sharp parts of structural elements);

- the danger of radiation (the effect of sound (acoustic), high-frequency, infrared, ultraviolet and ionizing radiation, as well as visible light of coherent high intensity (laser radiation);

- chemical danger (contact with some chemicals that are used to maintain the equipment or from inhaling their vapors).

The main danger that can arise during work is electric shock [26], as work is carried out with a personal computer and pairs of cameras connected to it. Although electrical accidents in the home (such as touching electrical outlets or being electrocuted by a small appliance) rarely result in significant damage or consequences, there are about 300 high-voltage electrical-related deaths in the United States each year. More than 30,000 non-fatal electrocutions occur annually in the United States, and electrical burns account for about 5% of admissions to burn units in the United States.

Traditionally, the severity of electrocution depends on Kouwenhoven factors:

- Type of current (direct or alternating current);
- Voltage and power (measure current strength);
- Duration of exposure (the longer the contact, the more severe the injury);
- Resistance of the body;
- Current path (determines damaged tissue).

According to the new concept, the strength of the electric field more accurately predicts the severity of the injury.

Alternating current often changes direction; this type of current typically supplies electrical outlets in European homes. Direct current does not change its direction; this current is produced by batteries. Defibrillators and pacemakers usually deliver direct current. How alternating current affects the human body depends mainly on its frequency. Low-frequency alternating current (50–60 Hz) is

used in home networks. Since low-frequency alternating current causes strong muscle contraction (tetany), which leads to the impossibility of removing the hand from the current source and, accordingly, increases the duration of its impact, such a current can be 3-5 times more dangerous than high-frequency alternating currents more dangerous direct current of the same voltage and strength. Direct current usually produces a single convulsive contraction that often throws the victim away from the current source.

Both alternating current and direct current are characterized by a regularity: the higher the voltage (V) and the current (A), the more serious the electrocution (with the same duration of exposure). Current in homes ranges from 110 V (standard electrical outlet) to 220 V. High voltage current causes deep burns, and low voltage current (110-220 V) causes muscle tetany and "freezing" to the current source. The maximum current strength, which can not only cause contraction of the flexor muscles of the hand, but also allow the hand to release the current source, is called the "releasing" current. The size of the "releasing current" depends on the muscle mass of the person. For a 70 kg male, the discharge current will be 75 milliamperes (mA) for AC and about 15 mA for DC.

A low-voltage alternating current with a frequency of 60 Hz passed through the chest for even a fraction of a second can cause ventricular fibrillation at a current as low as 60-100 mA; about 300-500 mA is required for direct current. If the current is applied directly to the heart (eg, through a cardiac catheter or pacemaker electrodes), $< 1\ mA$ of AC or DC current can cause ventricular fibrillation.

Damage to tissues under the influence of electric current is primarily caused by the conversion of electrical energy into thermal energy, which leads to thermal damage. Thus, for any given current strength and duration of exposure, the tissue with the highest resistance can be the most severely damaged. The resistance of the body (measured in Ohm/cm^2) is provided primarily by the skin, because all

internal tissues (except bones) have negligible resistance. The thickness of the skin and its dryness increase resistance; dry, well-keratinized intact skin has an average value of 20,000-30,000 Ohm/cm^2. In a calloused palm or foot, the resistance can reach 2-3 million Ohm/cm^2; while thin skin has a resistance of about 500 Ohm/cm^2. Resistance for damaged skin (for example, a cut, a wound, a needle puncture) or moist mucous membranes can be no higher than 200-300 Ohm/cm^2.

If the skin resistance is high, more electrical energy can be dissipated through it, resulting in more skin burns but less damage to internal organs. If there is little skin resistance, then skin burns are less or absent, and a greater amount of electrical energy is transmitted to internal structures. Thus, the absence of external burns does not exclude electrical injury, and the severity of external burns does not determine the severity of electrical injury.

Damage to internal tissues depends on their resistance as well as the density of the electric current (current per unit area; energy is concentrated when the same current passes through a smaller area). For example, when electrical energy enters through the hand (primarily through low-resistance tissues, such as muscles, vessels, nerves), the electric current density increases in the joints due to a significant proportion of the joint's cross-sectional area consisting of high-resistance tissues (for example, bone, tendon), which reduces the area of tissues with low resistance; thus, damage to tissues with low resistance is more severe in joints.

The path of the current through the victim's body determines which structures are damaged. Since alternating current reverses direction, the commonly used designations of "input" and "output" are inadequate; the terms "source" and "land" are more accurate. The hand is the most typical "source", followed by the head. The feet are the most typical "earth" point. An arm-to-arm or arm-to-leg current usually travels through the heart and can cause an arrhythmia. This current

path is more dangerous compared to leg-to-leg current. Current passing through the head area can cause damage to the central nervous system.

Electric field strength is the strength of electricity in the area where it is applied. Along with the Kouwenhoven factor, it also determines the degree of tissue damage. For example, passing a current of 20,000 ($20\ kV$) through the body of a man 2 m tall will produce an electric field with a strength of about 10 kV/m. In the same way, a current of 110 V, which passed through 1 cm (for example, through the lip of a child), will create an electric field of 11 kV/m; that is why a low-voltage current passing through a small volume of tissue can cause the same severe damage as a high-voltage current passing through a large volume of tissue. Conversely, if the voltage is considered primarily, rather than the strength of the electric field, small or minor injuries may be classified as high voltage injuries. For example, an electric shock received in winter from shuffling a foot on a carpet corresponds to a voltage of thousands of volts, but causes an insignificant injury.

Exposure to an electric field can cause cell membrane damage (electroporation) even when the energy is insufficient to cause any thermal damage.

Exposure to a low-voltage electric field causes an immediate discomfort (shock), but rarely serious or irreversible damage. Exposure to a high-voltage electric field causes thermal or electrochemical damage to internal tissues. Damage may include the following:

- Hemolysis;
- Protein coagulation;
- Coagulation necrosis of muscles and other tissues;
- Thrombosis;
- Dehydration;
- Detachment of muscle and tendon.

Damage caused by a high-voltage electric field can cause massive swelling, which, as blood clots in the veins and muscles swell, can lead to the development of compartment syndrome. Massive edema can cause hypovolemia and arterial hypotension. Muscle destruction can lead to rhabdomyolysis and myoglobinuria, as well as electrolyte disturbances. Myoglobinuria, hypovolemia, and hypotension increase the risk of developing acute kidney injury. The consequences of body dysfunction do not always correlate with the volume of destroyed tissue (for example, ventricular fibrillation can occur against the background of relatively small tissue destruction).

An employer who uses hired labor must ensure that their workplaces are comfortable and safe.

The size of one workplace should be at least 6 square meters. If necessary, adjacent workplaces of employees working with a computer should be separated by partitions up to 2 meters high. When determining the sufficient size of the room and workplace for one person, it is necessary to additionally take into account closets, safes, cabinets or other pieces of furniture or equipment that are in the room.

On the employee's desk, it is possible to place devices supporting work (printers, speakers, scanners), as well as places for storing documents, provided that this does not limit the visibility of the screen and does not interfere with the employee. In case of excessive noise or vibration of technical equipment, the employer must provide employees with anti-vibration mats.

The employee's work chair must be lifting and swiveling, easily adjustable in height and provide adequate support and a comfortable position for the person's back and spine.

It is prohibited when working with a PC:

- independently disassemble and repair the system unit (laptop case), monitor, keyboard, computer mouse, etc.;

- stick foreign objects into the ventilation holes of the PC, laptop or monitor;

- place metal objects, containers with water (vases, flower pots, glasses) on the PC system unit and peripheral devices, as a fire or electric shock may occur due to water entering the middle of the device.

If visual discomfort or other unpleasant sensations experienced, person should take a short break.

To reduce neuro-emotional tension, fatigue of the visual analyzer, improve cerebral blood circulation, overcome the adverse effects of hypodynamia, and prevent fatigue, it is advisable to perform a set of exercises during several breaks.

To connect the equipment to the electrical network, you must use the power cord supplied with the equipment; Do not use self-made electrical cords to connect to the network. To ensure electrical safety, device cords must be insulated with rubber. When turning on or off the remote control, you should hold the plug block; do not touch the electrical cord for this purpose.

To prevent electric shock to a person in the event of a short circuit, the body of the video surveillance control panel must be grounded. To avoid electric shock, do not remove the covers from the equipment during operation. Do not touch live parts of electrical equipment with your hands. You cannot transfer your work on the remote control to employees who do not have the right to work on this equipment.

If any malfunctions are detected in the operation of the video surveillance panel (for example, with the appearance of voltage on non-current parts of the structure), it is necessary to stop work, turn off the equipment, hang a warning sign and notify the immediate supervisor about it to organize repairs.

To avoid accidents, the employee should not troubleshoot the equipment himself. If electrical equipment ignites, it is necessary to immediately report the fire, disconnect the equipment from the electrical network and start extinguishing the fire with the help of primary fire extinguishing means; at the same time, it should be remembered that the use of fire extinguishing agents that conduct electric current (for example, water, foam) can lead to electric shock; therefore, powder or carbon dioxide fire extinguishers should be used to extinguish burning electrical equipment.

In the event of an accident, poisoning, or sudden illness, it is necessary to immediately provide first aid to the victim, call a doctor or help take the victim to the doctor, and then inform the supervisor about the event.

Security requirements after completion of work include, first of all, saving the information obtained after processing the video of the monitored system. Then turn off the PC, monitor or laptop. disable the stabilizer if the computer is connected to the network through it. Clean the workplace and the presence of damage on working internal and external devices.

5.4. Fire safety of industrial premises

Any design office, thermal power station, or enterprise can serve as an example for using the developed model. For example, this system will be used at Antonov State Enterprise. The plant is divided into several separate facilities: the Design Office, the aircraft design workshop, the aircraft technical inspection workshop and the aircraft software and aircraft instrument update workshop. Also, on the territory of the Antonov State Enterprise, there is its own fire department. Each of the departments has several fire extinguishers of various types, as well as a fire extinguishing system. The entrance to the territory of the enterprise is closely guarded, each department has at least two exits from its building.

A fire at this object can occur under the influence of electrical and non-electrical causes [27].

Non-electrical causes:

- violation of fire safety when using lighting and working tools (soldering iron, drill, industrial hair dryer, etc.);

- negligent handling of open sources of fire (smoking, use of a lighter, etc.);

- self-ignition and self-ignition.

Electrical causes:

- overload, spark from a short circuit, these factors will lead to heating of the conductors to ignition temperature;

- unsatisfactory contacts at the junctions of conductors and their strong heating as a result of high transient resistance during the flow of electric current;

- spark from static discharge;

- malfunctions in the windings of transformers, which will lead to an increase in temperature and further ignition.

The head of the enterprise, developing organizational measures to ensure fire safety in accordance with the Fire Safety Rules in Ukraine, appoints responsible persons for fire safety of individual territories, areas, zones, objects, structures, buildings and their parts, technological and engineering equipment, as well as for maintenance and operation of fire protection equipment.

Responsible persons are appointed by order of the enterprise. Persons must be familiarized with this order for signature. Each room must also have a sign with the name, first name and patronymic and telephone number of the person responsible for fire safety in a particular room.

The duties of those responsible for the fire prevention condition and operation of fire protection equipment are provided for in the job instructions,

duties, and regulations on the company's subdivision. Persons responsible for fire safety develop fire safety instructions and ensure the maintenance of the fire prevention regime in the relevant territory, site, structure, building and premises.

Those responsible for fire safety must be familiar with a number of requirements:

- Creating conditions for safe work;

- Minimizing the risk of fires;

- Provision of technical means of preventing the occurrence and extinguishing of fires, as well as minimizing their consequences;

- Control of compliance with fire protection requirements and legislation.

- Development and implementation of regulations on extinguishing fires, evacuation and rescue from places of fire and smoke of people and property (material values);

- Conducting briefings and training on fire safety issues for officials and employees engaged in works with increased fire danger.

5.5. Measures for protection from electric shock

The use of a computer and electrical devices necessary for work on non-current-carrying elements of the test stand design, as well as damage to the insulation of the conductive elements of the device can lead to a short circuit through the human body.

According to УE-2007 "Rules for the installation of electrical installations", all electrical installations are divided into two classes: with a voltage up to 1000 V and with a voltage above 1000 V. Measures to ensure electrical safety are developed, first of all, based on which of these classes the designed electrical installation.

Protection against damage when touching parts that are not normally located, but can be energized in the event of an insulation breakdown, is usually provided by protective grounding, zeroing.

Calculation of the grounding device.

Accordance with the requirements ГОСТ 12.1.030-81 "Electric safety. Protective grounding. Zeroing" is the initial data for the calculation of the grounding device:

- resistance of the grounding device, $R_{доп} = 4\ Ohm$;
- soil resistivity, $\rho = 20\ Ohm \cdot m$;
- increasing coefficient of seasonality, taking into account soil resistance $K_п'' = 2$ (for the grounding line) and $K_п' = 1.2$ (for vertical grounding electrodes);
- type of grounding device - corner, its length $l = 2{,}5\ m$, the width of the sides $b = 0{,}05\ m$, located at a depth of $0.8\ m$ below ground level;
- the type of grounding line - strip, its width $b = 0{,}04\ m$, laying depth $h = 0{,}8\ m$;
- the type of placement of vertical grounding devices - in a row, the distance between them:

$$a = 5 \cdot \frac{a}{l} = 2\ m.$$

The calculation is performed in the following sequence:

1. Determine the calculated value of the specific resistance of the soil for vertical grounding devices ($\rho_{разр}'$):

$$\rho_{разр}' = \rho \cdot K_п' = 24\ Ohm \cdot m,$$

2. Calculate the resistance to current spreading for a single corner grounding device:

$$R = \frac{\rho'_{\text{разр}}}{2 \cdot \pi \cdot l} \cdot \frac{4.2 \cdot l}{b} = \frac{24}{2 \cdot 3.14 \cdot 2.5} \cdot \frac{4.2 \cdot 2.5}{\ln 0.05} = 8.17 \; Ohm.$$

3. Determine the conditional number of vertical grounding devices according to the formula:

$$n' = R \cdot \frac{t}{R_{\text{доп}}} = 2.04 \; pieces.$$

4. Depending on the conditional number of grounding conductors (n'), the ratio and type of placement of vertical grounding conductors (in a row), we find the coefficient of use of corner grounding conductors ($\eta_{\text{эт}}$), which takes into account the mutual shielding of the corners: $\eta_{\text{эт}} = 0{,}77$.

5. Determine the actual specified number of vertical grounding devices according to the formula:

$$n = \frac{n'}{\eta_{\text{эт}}} = 3.$$

6. Calculate the length of the main line connecting the vertical grounding devices according to the formula:

$$L = 1.05 \cdot a \cdot n = 15.75 \; m.$$

7. Determine the resistance to the current flow for the grounding line by the formula:

$$R_m = \frac{\rho'_{\text{разр}}}{2 \cdot \pi \cdot l} \cdot \frac{2 \cdot L^2}{\ln b \cdot h} = \frac{40}{2 \cdot 3.14 \cdot 15.75} \cdot \frac{15.75^2}{\ln 0.04 \cdot 0.8} = 3.6 \; \text{Ом};$$

where L is the length of the strip, m; b is the width of the strip, m; d – conductor diameter, m; h is the depth of laying the main line, m.

8. Depending on the type of placement of vertical grounding devices, determine the utilization ratio $\eta_{\text{эт}}$ of the grounding main line (connecting strip),

which takes into account the mutual shielding of the main line and vertical grounding devices $\eta_{эm} = 0.9$.

9. Calculate the resistance to the spread of the entire grounded device, according to the formula:

$$R = \frac{R_{\text{М}} \cdot R_T}{R_T \cdot \eta_{эm} + R_m \cdot \eta_{эm} \cdot n} = \frac{3.6 \cdot 8.17}{8.17 \cdot 0.9 + 3.6 \cdot 0.77 \cdot 3} = 2.9 \; Ohm.$$

Next, in accordance with the National Electrical Code (NEC), select grounding conductors and the method of connecting conductors and grounding conductors.

10. Calculation of the cross-section of the grounding conductor. The cross-section of the grounding conductor is calculated using the formula:

$$S \geq \frac{I_{\text{К3}}\sqrt{t_{cp}}}{272} \geq 0.05 \; mm^2,$$

where S is the cross section of the grounding conductor, mm^2; I - short-circuit current (for electrical installations up to 1000 V, $I_{\text{К3}} = 125/Rз = 65.8А$, $Rз = 2.9$ Ом); $T_{\text{ср}}$ – protection activation time ($t_{\text{ср}} = 0.1 \; sec$).

According to ГОСТ 12.1.038-82 "Electrical safety. Maximum allowable values of touch voltage and currents" electrical installations are connected to the internal main line using copper and aluminum conductors with a cross section of 4-6 mm. Laying of grounding conductors made of strip steel is recommended to be carried out openly along the structure of the house, in order to make their maintenance more accessible.

Connection of the grounding wire to parts of the equipment should be made by welding or bolted connection. It should be available for inspection. For a bolted connection, measures should be taken to prevent corrosion of the connection.

Means of protection against static electricity. Static electricity that occurs at all stages of the production process is a harmful factor. The charge of static electricity leads to the formation of sparks, which can cause a flash, failure of integrated circuits.

The danger is represented by spark discharges from dielectrics, as well as from insulated metal surfaces, which are the cause of ignition and explosion of mixtures of air with dust, gas and steam. The ignition effect depends on the energy released during the discharge of static electricity. $W = 0.5 \cdot C \cdot U$, where: C is the system capacity; U is the potential of the system. As a result of the induction effect when a person moves, rubs against clothes, covering tables, chairs, panels - service personnel can be charged up to the potential $U = 40\ kV$.

The effect of electrostatic discharge becomes perceptible to a person at $U > 3\ kV$. A stable spark can occur already at $5\ kV$, at a potential discharge of up to $25\ kV$, a person has a mild convulsion, and at a potential of more than $35\ kV$ - an acute convulsion.

To estimate the value of the potential on the cotton clothing of a person working with devices, an experimental expression was obtained: $U = 2.5 - 0.049 \cdot \zeta$, where ζ is the humidity of the environment; The classification of the main means of protection against static electricity is given in ГОСТ 12.4.124-83 "Means of protection against static electricity. General technical requirements".

To eliminate the causes of the formation of static charge, it is necessary to use conductive materials to cover floors, panels, work tables, chairs, as well as shoes and clothes of service personnel.

To reduce the degree of electrification and increase the conductivity of dielectric surfaces, it is necessary to maintain the relative humidity of the air at the level of the maximum permissible value according to ГОСТ 12.1.005-88 "General sanitary and hygienic requirements for the air of the working area". When the shoes

are in contact with the floor, it is necessary to ensure a continuous flow of charge. This can be achieved if the resistance of the covering and shoes does not exceed 106 Ohm. At workplaces, all metal and electrically conductive non-metallic parts of technological, measuring and maintenance equipment are grounded. When carrying out maintenance and repair of the designed device, the worker performing the work must stand on a dielectric mat and use a tool with insulated handles.

5.6. Conclusion

This section discusses the dangerous and harmful factors that arise during the installation of a tracking system using stereo vision. Technical measures that eliminate or reduce the effect of harmful factors in the installation process were given. The calculation of lighting and zeroing of the operator's workplace was also carried out. Using this data, it is possible to ensure safe and non-harmful working conditions for workers.

CHAPTER 6. ENVIRONMENTAL PROTECTION

6.1. Current issues of atmospheric air pollution

Atmospheric air pollution [28] is one of the most serious environmental factors that affect the health of any person in low-, middle-, and high-income countries.

In 2016, worldwide, ambient air (outdoor air) pollution in both urban and rural areas was estimated to be responsible for 4.2 million premature deaths per year; this mortality was due to exposure to fine suspended particles up to 2.5 μm in diameter (PM2.5), which cause cardiovascular and respiratory diseases, as well as cancer.

Populations in low- and middle-income countries are disproportionately exposed to air pollution, and of the 4.2 million premature deaths, 91% occur in low- and middle-income countries, mainly in the WHO regions for South-East Asia and the West parts of the Pacific Ocean. Recent estimates of the severity of disease indicate that air pollution plays a very significant role in cardiovascular disease morbidity and mortality. There is a growing body of evidence, including from studies conducted in areas with high levels of air pollution, that suggest a link between air pollution and the risk of cardiovascular disease.

According to World Health Organization (WHO) estimates, in 2016 about 58% of premature deaths related to outdoor air pollution were caused by coronary heart disease and stroke, 18% each by chronic obstructive pulmonary disease and acute lower respiratory tract infections, respectively, and 6% lungs.

In some cases, death may be associated with several risk factors. For example, both smoking and air pollution are factors in lung cancer. Some lung cancer deaths could be prevented by improving air quality or reducing smoking.

In 2013, the WHO International Agency for Research on Cancer (IARC) concluded that atmospheric air pollution has a carcinogenic effect on human health, and air pollution by suspended particles is most associated with an increased incidence of oncological diseases, especially lung cancer. A link between air pollution and increased incidence of urinary tract/bladder cancer has also been established.

One of the most important tasks in the field of health care is the correction of all factors of the development of non-infectious diseases, including air pollution.

Most sources of outdoor air pollution are beyond people's control, and reducing their impact requires concerted action by local, national and regional authorities in sectors such as transport, energy, waste disposal, urban planning and agriculture.

Internal combustion engines negatively affect the health of millions of people every day. People suffer from cases of coughing, asthma attacks, acute and chronic bronchitis, as well as diseases of the heart and circulatory system. Professional drivers, car service workers, road workers, etc. are at the highest risk.

The WHO made a statement in 1998 about the connection between carcinogenic substances found in the exhaust of cars running on diesel fuel and lung cancer. A new type of substances found in the exhaust of cars running on diesel fuel, according to scientists, causes mutations in the human body and are the main factors that cause lung cancer.

The chemical composition of exhaust gases is so dangerous that it harms not only human and animal health, but also destroys trees and even houses.

Human health is determined not only by heredity and lifestyle, but also by the quality of the environment. Today, the deterioration of public health from an ecological point of view is primarily a consequence of atmospheric air pollution. The main sources of atmospheric pollution are road transport and industrial zones.

The sensitivity of the population to the effects of pollution depends on a large number of factors, including age, condition, general state of health, temperature, humidity, etc. The elderly, children, patients, smokers who suffer from chronic bronchitis, coronary insufficiency, asthma are considered more vulnerable.

Among the factors of direct action (all except environmental pollution), air pollution takes, of course, the first place, since air is a product of continuous consumption by the body.

The human respiratory system has a number of mechanisms that help protect the body from the effects of air pollutants. Hairs in the nose filter large particles. The sticky mucosa in the upper respiratory tract traps small particles and dissolves some gaseous pollutants. The mechanism of involuntary sneezing and coughing removes polluted air and mucus when the respiratory system is irritated.

Fine particles pose the greatest danger to human health, as they are able to pass through the natural protective membrane in the lungs. Inhaling ozone causes coughing, shortness of breath, damages lung tissue and weakens the immune system.

The state of health of the population is to some extent an indicator of the ecological situation. The most sensitive are the immune and endocrine systems, under whose control, along with the nervous system, the body as a whole and its individual structures are located.

The impact of atmospheric pollution on health can be divided into two types depending on the time of manifestation of the effect:

- acute, which occurs during or immediately after an increase in the concentration of a toxic substance;

- chronic influence, the result of which is not immediately apparent, but after a few hours, sometimes years.

6.2. Methods of improving the ecological state of atmospheric pollution

There are many examples of successful policies [29] in the fields of transport, urban planning, power generation and industry aimed at reducing air pollution:

- industry: implementation of environmentally friendly technologies that allow reducing emissions into the atmosphere at industrial enterprises; improvement of household and agricultural waste disposal methods, including capture of methane generated at waste disposal facilities as an alternative to burning it (for its further use as biogas);

- energy: ensuring access to inexpensive sources of energy in the home for cooking, heating and lighting;

- transport: transition to ecologically clean methods of electricity production; priority development of networks of high-speed urban transport, pedestrian and bicycle traffic in cities, as well as rail intercity freight and passenger transportation; the transition to cleaner diesel engines for heavy-duty vehicles, low-emission vehicles, and cleaner fuels, including low-sulphur fuels;

- urban planning: increasing the energy efficiency of buildings, greening and reducing the area of cities, which increase their energy efficiency;

- electricity production: increasing the use of fuels with low levels of emissions into the atmosphere and methods of electricity production that are not based on burning fuel (for example, solar, wind or hydropower); combined heat and electricity generation; distributed energy (for example, small-scale power supply networks and placement of solar panels on the roofs of buildings);

- utilization of household and agricultural waste: implementation of measures to reduce volumes, separation, secondary processing and secondary use of waste; improvement of methods of biological utilization of waste, such as anaerobic fermentation with production of biogas (a low-cost and feasible alternative to open burning of solid waste).

6.3. Types and impact of anthropogenic pollution on the atmosphere

Atmospheric pollution is changes in the gaseous composition of the atmosphere as a result of the content of impurities in it. Analysis of air pollution is perhaps the most difficult task of analytical chemistry, since air is a moving system, the composition of which is constantly changing, and one sample can contain dozens and hundreds of organic and inorganic compounds. In addition, the concentration of toxic substances in the atmosphere can be extremely low – 10.4 - 10.7% and below. Atmospheric pollution causes such global problems [30] as climate warming (greenhouse effect), acid rain, destruction of the ozone layer, desertification, and others. Natural sources of pollution are volcanic eruptions, dust storms, space dust, geysers, forest and steppe fires, excretion of plants, animals and microorganisms. Some harmful impurities (carbon dioxide, carbon monoxide, hydrogen sulfide, chlorides, and other compounds) enter the atmosphere from the ocean. The peculiarity of natural pollutants is that they are not as toxic as anthropogenic ones, and they act periodically. As a result of anthropogenic activity, artificial pollution is carried out, while the composition of atmospheric air changes significantly.

The following types of anthropogenic pollutants are distinguished according to the features of their structure and impact:

- mechanical - emissions from cement plants, smoke, soot, etc.;
- chemical - gaseous substances capable of reacting with air components, for example CO, CO_2, SO_2, NO_x.

Production and household processes play the most significant role in anthropogenic atmospheric pollution.

The main sources of anthropogenic pollution of the atmosphere in Ukraine and most countries of the world are:

- thermal power facilities (thermal power plants, furnaces of boiler houses and buildings);

- industrial enterprises (ferrous and non-ferrous metallurgy, chemical industry, mechanical engineering, objects of extraction and processing of natural minerals);

- transport (mainly motor vehicles);

- enterprises of the agro-industrial complex (livestock complexes, arable and crop production, canneries, etc.);

- construction sites.

In different countries, depending on the state of economic development, a different structure of the main sources of atmospheric pollution is noted. In the USA, due to the high degree of motorization, the contribution of transport to total emissions is about 60%, industry - 17%, energy - 14%, heating and waste disposal - 9%. In France, industrial pollution accounts for 37%, heating – 25%, transport – 23%, energy – 15%. In Ukraine, vehicles annually emit up to 2 million tons of harmful substances into the atmosphere, which is 31% of total pollution, including 63% of lead, 54% of carbon monoxide, and 25% of nitrogen oxide from the total emissions of these compounds.

More than 4 million tons of pollutants are emitted from stationary sources per year, which is about 64% of the total volume of pollutants. Almost 90% of the total volume of pollutant emissions into the atmospheric air by stationary sources in Ukraine is created by thermal power, ferrous metallurgy, coal, oil and gas mining, and cement industries. In terms of sectors of the economy, the largest share of emissions of pollutants - 41.3% (excluding carbon dioxide) falls on the production and distribution of electricity, gas, water, in particular, powerful thermal power plants (TPPs) that operate on low-grade, dusty fuel (as a result of burning of fuel, ash, soot, various gaseous products are emitted into the air), also boiler installations of thermal power plants produce a large amount of gross

emissions of substances that pollute the atmosphere (the main harmful impurities here are oxides of sulfur and nitrogen, dust and soot).

During 2015, the State Hydrometeorological Service monitored atmospheric air pollution in 53 cities of Ukraine at 163 stationary, two route observation posts and two cross-border transfer stations. The content of 30 pollutants was determined in the atmospheric air.

6.4. Pollution assessment methods

To assess air pollution [31], use:

- laboratory methods (characterized by high accuracy and indispensable for in-depth research);

- non-laboratory methods - express and automatic methods (ensure continuous monitoring of atmospheric air pollution).

Determination of concentrations of most pollutants in atmospheric air is usually performed by laboratory methods. During the laboratory eco-analytical control of hazardous substances in the air of background areas, settlements and in industrial emissions, control technology with separate procedures for selection and measurement of sample indicators is mainly used.

Most often, air pollution control uses laboratory analysis procedures using standardized as well as other official measurement methods (MM).

According to the established tradition in domestic gas analytical instrument construction and monitoring of atmospheric pollution, the means of controlling gaseous environments are divided into systems (complexes), devices and other technical means of air pollution control (TMPC), while grouping them according to the characteristics of the air environment being analyzed, as follows:

- TMPC of the atmosphere (atmospheric air);
- TMPC air of populated areas and residential premises;

- TMPC air of the working area and production premises;
- TMPC of emissions and vapor-air mixtures entering the atmosphere.

In the eco-analytical control of air pollution, the first three types of means are most often used, while the fourth group rather refers to the technical control of production processes and the control of emissions into the atmosphere. However, since these tools allow characterizing the source of environmental impact (pollution), it also makes it possible to classify them as eco-analytical control tools - namely, environmental impact monitoring.

At the same time, TMPC atmospheres are also traditionally divided according to the degree of their automation:

- on automatic gas analyzers (means of measuring the content of hazardous substances, as well as determining the qualitative and quantitative composition of the gas mixture) and/or gas alarms (means of indicating the level of pollution);
- non-automated devices or other means of controlling vapor-air mixtures and gas environment (for example, manual express gas detectors).

On the modern Ukrainian market of eco-analytical control tools, automatic gas detectors are the most common, which are gradually being replaced by gas analyzers, which are measuring devices. A significant share of this market is also occupied by non-automated TMPC atmospheres based on indicator tubes, semi-quantitative express tests, indicator tickets, pieces of paper, paints, chalk, films, etc.

Such means are the most accessible for public ecological control and educational ecomonitoring due to their simplicity and cheapness. Professionally, they are most often used during field control "on the spot" - to solve the first task in the technological chain - "research", as well as for preliminary and very approximate ("semi-quantitative") measurements.

6.5. Environmental and economic damage assessment

The ecological and economic assessment [32] of damage caused to the surrounding natural environment consists in determining the actual and possible material and financial losses and damages from the deterioration due to anthropogenic influence of the qualitative and quantitative parameters of the natural environment as a whole and its individual ecological and resource components (water) resources, land resources, resources of the plant and animal world.

Environmental damage is estimated as the costs of society associated with a change in quantitative and qualitative indicators characterizing the state of the environment, and is composed as follows:

- Additional costs of society in connection with various changes in the environment;
- Costs for returning the environment to its former state;
- Additional costs of the future society in connection with the irreversible removal of a part of scarce natural resources.

A mandatory environmental requirement is to maintain the quality of the natural environment: atmosphere, hydrosphere, lithosphere. The maximum permissible concentrations of harmful substances in the air of residential buildings are set for 200 substances, for the working zone they are standardized for 703 substances, in water objects for economic, drinking and cultural and household purposes, there are 640 names of ingredients, 140 names are included in the list of pollutants in the soil.

Environmental protection measures are quite diverse and can be combined into large groups according to their purpose.

The first group includes single-purpose protective measures, the sole purpose of which is the complete exclusion or reduction of industrial pollution of the environment.

The environmental protection effect of these measures is determined by the installation of standard environmental protection equipment for the enterprise, the development and implementation of new, more effective methods of water and gas treatment, the introduction of certain changes in the technology of manufacturing products, which lead to a decrease in the concentration of harmful components in industrial effluents and gas emissions, etc.

The second group combines single-purpose resource-saving scientific research, which is carried out with the aim of saving raw materials, fuel and energy. The resource-saving effect of developments of this kind is obtained as a result of reducing the consumption rates of raw materials and fuel and energy resources in the field of production, reducing losses during transportation and storage of labor items, using new less material and energy-intensive technological processes and productions, etc.

The measures of the third group are multi-purpose. Environmental protection tasks are solved here along with the problems of improving product quality, improving the use of raw materials, fuel, energy, technological equipment, labor and other elements of material production.

The structure of the economic effect obtained during the activities of the third group is ambiguous depending on their type.

Technical and economic substantiation of environmental protection measures in the general case includes:

- analysis of production, scientific and technical, techno-economic, social and environmental problems in the field related to the analyzed research; a brief description of the goals and content of the work under consideration, as well as the

expected, planned or actual technical and environmental characteristics of the environmental protection equipment or new technology;

- assessment of research deadlines and funding amounts;

- selection of the comparison base; analysis of technical and environmental advantages of new technology compared to the basic one;

- identification of possible sources of economic effect at the facility where implementation is carried out and in related industries at production;

- establishing the scope of implementation of the development;

- determination of costs for the implementation of a basic protective measure;

- calculation of the amount of economic damage from environmental pollution when implementing a basic protective measure;

- determination of costs for the development and implementation of proposed options for environmental measures;

- calculation of the amount of economic damage from environmental pollution during the implementation of the proposed environmental protection measures;

- bringing options to a comparative form;

- economic comparison of options; determining the size of the economic effect;

- assessment of the socio-economic effect of the application of development results; evaluation of the scientific and technical effect of the conclusions.

Damages can be instantaneous, permanent, latent (appear over time), etc. In addition, economic losses are divided into potential and calculated.

Potential damage is economic damage, for the elimination of which no additional costs are currently required.

Estimated damage is that part of damage that manifests itself in a certain period and can be expressed in monetary form at a given level of development of

economic science. As knowledge grows, the estimated damage will tend to the potential, and the latter to the actual damage.

There are two methodological approaches to determining the economic damage caused by pollution:

- indirect (consolidated by spheres of influence);
- recipient (on the basis of a direct account by detailed elements of influence).

The detailed calculation is based on the data of the analogue object, actual statistical materials, expert assessments.

The following damage structure is usually distinguished:

1) public health;
2) communal and household management;
3) agriculture and forestry;
4) transport, industry, etc.

Each of the constituent elements (local damages) is determined separately, and then they are all summed up. This approach to determining damage is called the method of summing up local damage methods or direct calculation. In this case, the calculation is carried out in the following sequence:

1) determination of emissions (discharges) of harmful impurities from the sources of their formation;
2) establishing the concentration of impurities in the atmosphere (water body);
3) definition of natural damage and economic damage.

Such a conceptually simple scheme for determining damage is associated with great difficulties, because it requires detailed information about the change in physical characteristics. A special study should be conducted in each case.

In practice, they usually use the indirect method of aggregated assessment of economic damage - the method of calculation based on the "monopollutant", which gives an approximate estimate, but can be a reference point for solving joint tasks.

An indirect approach to the assessment of economic damage involves the use of a number of indicators that reflect the value of damages, factors, the product of which allows determining the consolidated amount of economic damage. In the consolidated calculation, three groups of spheres of influence are distinguished - atmosphere, water, land, which are state and branch consolidated estimates of specific damage.

The approach based on a simplified procedure is reduced to a single formula calculation. The idea behind the calculations is that initially all harmful impurities emitted into the atmosphere or discharged into water bodies lead to a "monopollutant". Of course, 5 tons of lead and 3 tons of nitrogen oxide cannot be stored directly. Their impact on the environment and people is different. However, if assume that it is possible to know how many times one pollutant is more dangerous than another, then it is possible to assign weighting factors A_i to each of them. After the volumetric indicators are multiplied by the weighting factors, they can be added together. Through the sum, we get the conditional mass of emissions $\sum A_i$, i.e. some conditional "monopollutant" that characterizes the general level of environmental pollution.

Then, this mass of emissions is multiplied by a coefficient, which is laid down by the features of a certain region. This coefficient allows you to take into account the reaction of a specific territory to emissions of harmful substances. In northern regions, where the ability of the environment to absorb harmful impurities is small, the coefficient is higher than where nature more easily copes with harmful effects.

6.6. Pollution by metallurgical enterprises

Experts believe that among all enterprises, the most pollution comes from ferrous and non-ferrous metallurgy plants. As a result of their activity, emissions of harmful substances enter the air. Metallurgical enterprises discharge [33] a large amount of wastewater into reservoirs. In order to reduce the negative impact of factories on the environment, they need treatment facilities. The old ones need to be replaced with new ones and used to their full potential.

Heat supply is a major branch of the modern economy, 25% of all fuel extracted and produced in Ukraine is spent annually on these needs. In conditions of limited fuel resources, their rational use is a task of great national importance. In addition, heating installations are sources of air pool pollution. In this regard, it is necessary to make a transition from small heat sources, with a low efficiency and a high concentration of emissions of toxic substances, to the use of large centralized heat sources, the gas emissions of which contain a smaller amount of toxic substances. Currently, there is a trend of increasing emissions in the industrialized regions of the country. In Ukraine, most of the industrial sectors, according to the classification of the International Agency for the Study of Cancer under the WHO, belong to the carcinogenic dangerous both for those who work for them and for the population, in general, because they are a source of education and emission into the surrounding natural environment of carcinogenic substances and their predecessors.

It is believed that thermal power plants (TPP) produce approximately 25% of harmful emissions entering the atmosphere from industrial enterprises. However, for example, according to data in the city of Kramatorsk, the contribution to air pollution from the operating TPP in 2011 was 70% or 6.5 thousand tons - despite the fact that this city has several large metallurgical and machine-building plants.

Adequate assessment of the real magnitude of economic damage from environmental violations is necessary in all sectors of the economy: for the purpose of assessing the overall economic efficiency of the economy, for the selection of investment projects of all kinds, in particular projects for environmental protection, the introduction of market management methods, the implementation of the principles of environmental insurance.

Recently, in connection with the wear and tear of heat supply systems, much attention has been paid to increasing the energy security of heat supply to the population. The problem of energy security should be solved by preventive preventive measures in order to exclude emergency situations at heat sources during the heating period and the associated large financial and social costs. This applies to all heat sources operating on any type of fuel - the difference is only in the amount of emissions into the atmosphere).

In 2010 The State Environmental Inspection of Ukraine collected about 42 million hryvnias in the form of fines for environmental pollution.

But today, the legislation does not clearly state how exactly the amount of emissions into the atmosphere should be calculated. Accordingly, enterprises calculate it according to different methods, as a result, the payments of two companies that pollute the air in the same way may differ by several times.

The degree of research into the problem of economic damage is characterized by the presence of specific developments in the field of assessing damage caused by an enterprise to the surrounding natural environment. Currently, in developed foreign countries, the economic damage caused by environmental pollution is estimated at 2-7% of the gross national product (GNP); for comparison, in the 70s of the 20th century, losses were estimated at 6-14 GNP. At the current stage, financial investments in environmental protection measures in countries with developed economies amount to 4-6% of GNP).

The general principle of economic behavior is the comparison of costs and results of economic activity. In the case of individual measures for environmental protection, costs are understood as funds directed to environmental protection measures, and results are economic losses that are prevented.

In the case of assessing the overall efficiency of all economic activity, the economic damage caused by this activity should be taken into account as one of the (negative) results. The lack of reliable estimates of economic damage leads to the fact that indicators of economic damage as a result of management are not taken into account at all, and in this connection, nature management falls out of the general scheme of evaluating the effectiveness of any economic activity.

Of course, this circumstance inhibits the implementation of all measures aimed at environmental protection.

It is the indicator of economic damage from environmental violations that allows you to transfer the negative impact of projects, plans and programs on the environment from the category of environmental characteristics of the project to the category of its economic characteristics.

Environmental damage refers to the losses that the natural environment itself suffers. However, in this case, it is more correct to talk about damage to the environment, and moving to the economic aspect of this problem, about the costs of eliminating damage to the environment caused by the activity of an industrial enterprise.

The task of eliminating such damage is set not only because of a humane attitude to nature: due to environmental violations, the economy suffers losses, since functioning in a disturbed environment significantly worsens the main technological parameters: material intensity, capital intensity, labor productivity, etc.

The determination of economic damage is based on the value expression of environmental quality losses and ecological damage.

The main tasks of assessing economic damage are determined by existing social relations and specified depending on the real state of the natural environment. Generalizing tasks of economic damage assessment are:

1) The assessment of economic damage to the surrounding natural environment is necessary for inclusion in the ecological passport of an industrial enterprise, which characterizes the degree of influence of the enterprise and the suitability of the natural environment for meeting the needs of other subjects;

2) The application of economic damage assessment is necessary in order to compare the extent of economic damage caused by the enterprise to atmospheric air, water and land resources. This approach will allow zoning of the territory according to the amount of damage and will show the value of this region within the borders of Ukraine;

3) With the development of market relations, the task of assessing economic damage is to regulate the relations of an industrial enterprise with a natural object during the purchase, sale, privatization, lease, and crediting of the enterprise;

4) The task of assessing economic damage is to determine the planned and actual indicators that correspond to the limit and actual volume of pollutant emissions into the natural environment.

6.7. Application of the developed system for estimating the volume of emissions

The developed design uses stereo vision based on observation from two web cameras. This model works identically to human vision, conditionally speaking, every person can estimate the conditional distance and size of a body that is at a distance of no more than 20 meters, since the idea of the size of the object may be wrong as the distance increases. Unlike human vision, the developed model can be adjusted in such a way to observe objects without paying attention to the difficulties

with increasing distances. This is done by increasing the distance between the optical lenses, which are directed parallel to the object of observation. That is, after proper calibration and preparation of the installation for work, it will be able to observe many voluminous and distant bodies with a guarantee of accuracy, the error of which does not exceed 3-5%.

Let it be given an example related to environmental pollution from the fuel energy industry. It is known that exhaust or emission gases are products of oxidation and incomplete combustion of hydrocarbon fuel, coal or other type of fuel, or are caused by the process of burning waste. Gas emissions are the main cause of exceeding permissible concentrations of toxic substances and carcinogens in the atmosphere of large cities, the formation of smog, which is a frequent cause of poisoning in confined spaces. The main goal of determining the impact of emissions of harmful substances is monitoring the level of emissions to the outside, that is, the process of gas and steam formations.

Let's choose a thermal power plant (Fig. 6.1), which emits a large amount of gases into the atmosphere. A thermal power plant (or a thermal power plant) is a power plant that produces electrical energy by converting the chemical energy of the fuel during combustion into thermal energy, and then into mechanical energy of rotation of the generator shaft. Various combustible fossils are widely used as fuel: coal, natural gas (propane, methane), less often - fuel oil, hydrogen, biogas, shale gas, oil, gasoline, diesel fuel, waste alcohol, peat, combustible shale, firewood.

Many large thermal plants produce only electricity - Condensing power plant; medium stations can be used to generate heat in heat supply schemes (CHP).

In traditional thermal power plants, fuel is burned in the furnace of a steam boiler, heating and turning into steam water that is pumped inside the boiler in special tubes (water-tube boiler). The resulting superheated steam with a high temperature (up to 400-650 degrees Celsius) and pressure (from units to tens of

MPa) is fed through a steam pipe to a turbogenerator — a combined steam turbine and electric generator. In a multi-stage steam turbine, the thermal energy of the steam is partially converted into the mechanical energy of rotation of the shaft on which the electric generator is installed. In CHP, part of the thermal energy of the steam is also used in network heaters.

In a number of thermal power plants, the gas turbine scheme has become widespread, in which the mixture of hot gases obtained during the combustion of gaseous or liquid fuel directly rotates the turbine of the gas turbine installation, the axis of which is connected to the electric generator. After the turbine, the gases remain hot enough to be usefully used in a recovery boiler to power a steam engine (steam plant) or to provide heat (gas turbine CHP). Energy is one of those sectors of the world economy in which changes are necessary to avoid the unacceptable consequences of global warming. Estimates of energy infrastructure based on the global CO_2 emission budget show that no new fossil fuel power plants should be commissioned after 2017. Solid waste of thermal power plants - ash and slag - is close to metallurgical slag in composition. Their output now amounts to about 70 million tons per year, and about half of this waste consists of ash from the burning of hard coal. The degree of use of ash and slag waste does not exceed 1.5-2%. By chemical composition, these wastes consist of 80-90% SiO_2, Al_2O_3, FeO, Fe_2O_3, CaO, MgO with significant fluctuations in their content. In addition, these wastes include the remains of unburned fuel particles (0.5-20%), compounds of titanium, vanadium, germanium, gallium, sulfur, and uranium. The chemical composition and properties of ash and slag waste determine the main areas of their use.

The main task will be to estimate the volume of gas that spreads over time.

The system will work according to the principle of streaming shooting, that is, the tracking process will take place every half a second. Let one or more pipes be chosen, which are responsible for the release of gases into the air, it becomes

impossible to track a larger number due to the inability to observe such a large amount of data.

Fig. 6.1. An example of emissions from a thermal power plant

The cameras will be directed to the final point of the pipeline, for more capture of the footage of the volume of gases. Streaming video capture tracks the cloud created by fuel processing and estimates the distance, volume and all parameters to determine the amount of the dispersant, so for environmental researchers the system outputs monitoring data of hazardous gas emissions into the atmosphere over time. This is a significant advantage of using this system, as the mechanism has a low cost and produces high accuracy results. In further processing, the data can be transferred to the WHO to check the functionality of the emission monitoring system.

This system is suitable for detecting this type of problem because it can be installed for a large number of pipelines that create environmental pollution problems.

Also, this system can be useful when monitoring environmentally dangerous objects, such as the Chernobyl Nuclear Power Plant or when working with the consequences of chemical, radioactive, bacteriological and other pollution as a result of the aggressive war of the Russian Federation on the territory of Ukraine.

The use of the modeling method provided by this system is extremely important for determining the level of environmental danger, the level of potential harm to a person in contact with such an object, and the geographical boundaries of these objects.

CONCLUSION

During investigation, it was established that a useful stereo pair could be made for assessing the surroundings and creating 3D landscapes utilizing the MATLAB software environment and two web cameras.

The method's adjustment, namely calibration, getting a good 3D model, etc., are its key drawbacks. Another challenge was building a solid platform that would allow the cameras to be moved around and their location adjusted in relation to the axes.

Additionally, subtle differences were seen between various areas, projective planes, and the graph itself. Because the light is not steady, some glare, body transparency, or vice versa, specularity, are possible. Mirrors, which were excluded from consideration since the cameras couldn't comprehend and match positions with the actual world and things in the mirror didn't have a spatial location, revealed an important fact.

Given that the setup is made of inexpensive equipment and materials, the result can be viewed as being quite successful. An assessment of the range accuracy revealed that the system does not reduce a relative error of over than 5%, when it is used indoors, which is a result that is generally acceptable even for rangefinder types of work using other methods.

Under the condition of tracking outdoor objects, the system showed that with increasing distances the coefficient of relative error increases, and the system becomes unusable, but this problem can be solved with proper adjustment of the structure itself and calibration with adjustment for long distances.

Based on studies from the point of view of ecology, it was found that the system may be suitable for measuring the amount of emissions of pollutants into the atmosphere, i.e. gases after combustion at thermal power plants and others.

Thus, it is a solution to an actual problem today, because this system can be used in real time.

As for labor protection, general methods were created for performers at the enterprise for comfortable and safe use of the device, and coefficients were also calculated for creating a grounding pole at the enterprise, to avoid accidents from electric shock.

REFERENCES

1. Где сегодня применяется компьютерное зрение. *Центр Международной Торговли*. URL: https://corp.wtcmoscow.ru/services/international-partnership/actual/gde-segodnya-primenyaetsya-kompyuternoe-zrenie/ (date of access: 27.08.2022).

2. Что такое компьютерное зрение и где его применяют | РБК Тренды. *РБКТренды*.URL: https://trends.rbc.ru/trends/industry/5f1f007e9a7947 56fafbfa83 (date of access: 28.08.2022).

3. Ализар А. Генерация 3D-моделей по фотографиям. *Все публикации подряд / Хабр*. URL: https://habr.com/ru/post/64080/ (date of access: 29.08.2022).

4. Махмутова Г.Э. Моделирование и исследование процессов стереовидения. МЭИ 2012.

5. Волосов Д. С. Фотографическая оптика. М., «Искусство», 1971.

6. Лидар. Применение технологии LiDAR. Карты и беспилотные автомобили | Gistroy. *Настольный лазерный гравер. Купить мини лазерный гравер с ЧПУ | Gistroy*. URL: https://gistroy.ru/article/lidar/ (date of access: 29.08.2022).

7. Датчик расстояния HC-SR04 - ультразвуковой модуль Ардуино. *ArduinoMaster все об Ардуино*. URL: https://arduinomaster.ru/datchiki-arduino/ultrazvukovoj-dalnomer-hc-sr04/ (date of access: 28.08.2022).

8. "Как устроены и работают инфракрасные датчики движения" Сайт для электриков - советы, примеры, схемы. *Сайт для электриков - советы, примеры, схемы*. URL: http://electrik.info/main/automation/917-kak-ustroeny-i-rabotayut-infrakrasnye-datchiki-dvizheniya.html (date of access: 30.08.2022).

9. Kinect.*ru.freejournal.org*. URL: https://ru.freejournal.org/1818924/1/kinect.html (date of access: 30.08.2022).

10. Как вычислить расстояние до объекта по фотографии. *Ретротехника, самоделки и борьба с идиотизмом – LiveJournal.* URL: https://bootsector.livejournal.com/43436.html (date of access: 27.08.2022).

11. Технология NanoLOC. *Самый информированный сервер микроэлектроника, описания - rs232, rs 232, микросхемы, hd44780, atmel, ацп, цап, irda, микроконтроллер.* URL: http://www.gaw.ru/html.cgi/txt/doc/Wireless/nanonet/index_loc.htm (date of access: 31.08.2022).

12. Точность определения расстояний с помощью технологии nanoLOC - Журнал Беспроводные технологии. *Журнал Беспроводные технологии.* URL: https://wireless-e.ru/rtls/receive-signal-strength-indication/ (date of access: 31.08.2022).

13. Бакулев П.А. Радиолокационные системы: *Учебник для вузов.: Радиотехника*, 2004.

14. Технология построения 3D-моделей объектов по набору изображений URL: http://masters.donntu.org/2012/fimm/solovyov/library/article7.htm (date of access: 31.08.2022).

15. LeCun, Y., Bottou, L., Bengio, Y. and Haffner, P. (1998), "Gradien-based learning applied to document recognition". *Proceedings of the IEEE*, Vol. 86, Issue 11, Nov. 1998, pp. 2278- 2324.

16. Krizhevsky, A., Sutskever, I. and Hinton, G.E. (2012), "ImageNet Classification with Deep Convolutional Neural Networks". *NIPS'12 Proceedings of the 25th International Conference on Neural Information Processing Systems*, pp. 1097-1105.

17. Карта глубины – Викиконспекты. *Northern Eurasia Contests.* URL: https://neerc.ifmo.ru/wiki/index.php?title=%D0%9A%D0%B0%D1%80%D1%82%D0%B0_%D0%B3%D0%BB%D1%83%D0%B1%D0%B8%D0%BD%D1%8B (date of access: 30.08.2022).

18. Kurakin A. Основы стереозрения. *Все публикации подряд / Хабр*. URL: https://habr.com/ru/post/130300/ (date of access: 01.09.2022).

19. Стереоскопическое компьютерное зрение URL: https://compress.ru/article.aspx?id=9529 (date of access: 03.09.2022).

20. Гонсалес Р., Вудс Р. Цифровая обработка изображений. – М.: Техносфера, 2005.

21. Buades, A., B. Coll, and J.M. Morel, "A review of image denoising algorithms, with a new one," *SIAM Multiscale Modeling and Simulation*, vol. 4, pp. 490.530, 2005.

22. Грузман И.С, В.С. Киричук, В.П. Косых, Г.И. Перетягин, А.А.Спектор. Цифровая обработка изображений в информационных системах: Учебное пособие. — Новосибирск: Изд-во НГТУ, 2000. — 168.

23. Апальков И.В., Хрящев В.В. Удаление шума из изображений на основе нелинейных алгоритмов с использованием ранговой статистики. — Ярославский государственный университет, 2007.

24. How to Describe the Multiple Types of Security Systems Present in Many Organizations - Birdseye Security Solutions. Birdseye Security Solutions. URL: https://birdseyesecurity.com/how-to-describe-the-multiple-types-of-security-systems-present-in-many-organizations/ (date of access: 03.10.2022).

25. Working environment | Health and safety | UNISON National. UNISON National. URL: https://www.unison.org.uk/get-help/knowledge/health-and-safety/working-environment/ (date of access: 08.10.2022).

26. Electrical injuries - Electrical safety at work. HSE: Information about health and safety at work. URL: https://www.hse.gov.uk/electricity/injuries.htm (date of access: 10.10.2022).

27. Fire Protection Tips for Industrial Buildings - Vanguard. Vanguard. URL: https://vanguard-fire.com/fire-protection-tips-for-industrial-buildings/ (date of access: 12.10.2022).

28. Air Pollution: Current and Future Challenges | *US EPA*. *US EPA*. URL: https://www.epa.gov/clean-air-act-overview/air-pollution-current-and-future-challenges (date of access: 14.10.2022).

29. 10 Best Ways to Reduce Air Pollution & Live With Clean Air. *AQI India*. URL: https://www.aqi.in/blog/10-best-ways-to-reduce-air-pollution/ (date of access: 15.10.2022).

30. Impacts of Anthropogenic Pollutants on Benthic Prokaryotic Communities in Mediterranean Touristic Ports. Frontiers. URL: https://www.frontiersin.org/articles/10.3389/fmicb.2020.01234/full (date of access: 16.10.2022).

31. Methods for Assessing Long-Term Exposures to Outdoor Air Pollutants - Current Environmental Health Reports. *SpringerLink*. URL: https://link.springer.com/article/10.1007/s40572-017-0169-5 (date of access: 17.10.2022).

32. Analyzing Health, Social and Economic *Effects under the Impact Assessment Act.* URL: https://www.canada.ca/en/impact-assessment-agency/services/policy-guidance/practitioners-guide-impact-assessment-act/analyzing-health-social-economic-effects-impact-assessment-act.html (date of access: 18.10.2022).

33. Влияние металлургии на окружающую среду: источники загрязнения, экологические проблемы и последствия. *Greenologia*. URL: https://greenologia.ru/eko-problemy/metallurgicheskay-promyshlennost.html (date of access: 19.10.2022).

APPENDIX

Appendix A. The program created for the location of the central axes on the camera image in the Matlab environment

```matlab
vid1 = videoinput('winvideo', 1, 'MJPG_1920x1080');
%second value is the port number (worth clarifying), the
third is the resolution (left camera)
vid2 = videoinput('winvideo', 3, 'MJPG_1920x1080');
%The right camera input

vid1.TriggerRepeat = Inf;   %Continuous image capture
vid2.TriggerRepeat = Inf;

preview(vid1);        % preview from cameras
preview(vid2);

 global LOOP_RUNNING;
LOOP_RUNNING = true;

while (LOOP_RUNNING)

    imgR = getsnapshot(vid2);
    imgL = getsnapshot(vid1);

    %creating centerlines

    subplot(1,2,1), imshow(imgL)
    hold on;
    line( [ 0, 1920 ], [540, 540 ], 'Color', 'red',
    'LineStyle','-');
    line( [960, 960], [0, 1080] , 'Color', 'red',
    'LineStyle','-');
    subplot(1,2,2), imshow(imgR)
    hold on;
    line( [ 0, 1920 ], [540, 540 ], 'Color', 'red',
    'LineStyle','-');
    line( [960, 960], [0, 1080] , 'Color', 'red',
    'LineStyle','-');

    pause(0.5);
end


stop (vid1);
stop (vid2);
```

Appendix B. Program designed to shoot 50 pairs of stereo images

```matlab
vid1 = videoinput('winvideo', 1, 'MJPG_640x360');
vid2 = videoinput('winvideo', 3, 'MJPG_640x360');

preview(vid1);
preview(vid2);

pause(5);

imgL = getsnapshot (vid1);
imgR = getsnapshot (vid2);

stop (vid1);
stop (vid2);

n=2;                    % pause time between shots
    tL_1=['E:\Left\'];        % save folders link
    tL_3 =['.jpg'];

    tR_1=['E:\Right\'];
    tR_3=['.jpg'];
for i = 1:50
    imgL = getsnapshot (vid1);
    imgR = getsnapshot(vid2);

    tR_2 = int2str(i);
    tL_2 = int2str(i);
    tL = strcat(tL_1,tL_2,tL_3);
    tR = strcat(tR_1,tR_2,tR_3);
    imwrite (imgR,tR ,'jpg')
    imwrite (imgL, tL, 'jpg')
    pause(n);

end
```

Appendix C. 3D scene output program

```matlab
showExtrinsics(stereoParams);


imageFileRight = 'E:\Images\imgLeft.jpg';
imageFileLeft = 'E:\Images\imgRight.jpg';


readerLeft = imread(imageFileLeft);
readerRight = imread(imageFileRight);


readerLeft = medfilt3(readerLeft);
readerRight = medfilt3(readerRight);


sigma = 1;
readerLeft = imgaussfilt3(readerLeft, sigma);
readerRight = imgaussfilt3(readerRight, sigma);

figure;
subplot(2,1,1)
imshowpair (readerLeft, readerRight);
title ('Unrectified Stereo Image');


[imgLeftRect, imgRightRect] = ...
    rectifyStereoImages(readerLeft, readerRight,
stereoParams,'OutputView','valid');


figure;
 imshow(stereoAnaglyph(imgLeftRect, imgRightRect));
 title('Rectified Image Frames');


imgLeftGray  = rgb2gray(imgLeftRect);
imgRightGray = rgb2gray(imgRightRect);

disparityMap = disparity(imgLeftGray, imgRightGray);

figure;
imshow(disparityMap, [0 64]);
title('Disparity Map');

figure;
colormap jet
colorbar
```

```matlab
points3D = reconstructScene(disparityMap, stereoParams);
points3D = points3D ./ 1000;
ptCloud = pointCloud(points3D, 'Color', imgLeftRect);

player3D = pcplayer([-1.5, 1.5], [-1.5, 1.5], [0, 7], 'VerticalAxis', 'y', ...
    'VerticalAxisDir', 'down');
view(player3D, ptCloud);
```