

Міністерство освіти і науки України
Харківський національний університет радіоелектроніки

Факультет інформаційних радіотехнологій та технічного захисту інформації
(повна назва)

Кафедра медіаінженерії та інформаційних радіоелектронних систем
(повна назва)

КВАЛІФІКАЦІЙНА РОБОТА
Пояснювальна записка

рівень вищої освіти другий (магістерський)

Використання методів фотограмметрії для визначення типу перешкод
мобільними роботами.
(тема)

Виконав:
студент 2 курсу, групи МІм-21-1
Осипенко А.П.
(прізвище, ініціали)

Спеціальність 172 Телекомунікації та
радіотехніка
(код і повна назва спеціальності)

Тип програми освітньо-професійна

Освітня програма Медіаінженерія
(повна назва освітньої програми)

Керівник доц. Горелов Д.Ю.
(посада, прізвище, ініціали)

Допускається до захисту

Зав. кафедри _____ Володимир КАРТАШОВ
(підпис)

2022 р.

Харківський національний університет радіоелектроніки

Факультет інформаційних радіотехнологій та технічного захисту інформації

Кафедра медіаінженерії та інформаційних радіоелектронних систем

Рівень вищої освіти другий (магістерський)

Спеціальність 172 Телекомунікації та радіотехніка
(код і повна назва)

Тип програми освітньо-професійна

Освітня програма Медіаінженерія
(повна назва)

ЗАТВЕРДЖУЮ:

Зав. кафедри _____
(підпис)

« ____ » _____ 20 ____ р.

ЗАВДАННЯ НА КВАЛІФІКАЦІЙНУ РОБОТУ

студентові Осипенку Андрію Павловичу
(прізвище, ім'я, по батькові)

1. Тема роботи Використання методів фотограмметрії для визначення типу перешкод мобільними роботами.

затверджена наказом по університету від " 24 " 10 2022 р. № 1384 Ст _____

2. Термін подання студентом роботи до екзаменаційної комісії 07.12.2022 р.

3. Вихідні дані до роботи Провести аналітичний огляд методів побудови карти глибини по стереозображенню. Дослідити відомі методи побудови карти глибини: методи ректифікації, методи пошуку пов'язаних точок, методи обробки карти глибини. Синтезувати комбінований метод побудови карти глибини з підвищеною точністю. Синтезований алгоритм реалізувати в середовищі MATLAB. Провести дослідження комбінованого методу з використанням тестових зображень. Методи дослідження – аналіз відомих алгоритмів, модельний експеримент, синтез комбінованого алгоритма, статистична обробка експериментальних даних.

4. Перелік питань, що потрібно опрацювати в роботі _____
Вступ

1. Аналітичний огляд алгоритмів розрахунку мапи глибини по зображенню стереопари.

2. Дослідження відомих алгоритмів побудови мапи глибини по зображенню стереопари.

3. Розробка і дослідження комбінованого алгоритму обчислення мапи глибини з підвищеною точністю.

Висновки

Перелік посилань

Додатки

5. Перелік графічного матеріалу із зазначенням обов'язкових креслеників, схем, плакатів, комп'ютерних ілюстрацій

1. Стереоскопічна система (1 аркуш А4).
2. Загальний алгоритм побудови карти глибини (1 аркуш А4).
3. Постановка задачі (1 аркуш А4).
4. Критерії відбору тестових стереопар (1 аркуш А4).
5. Алгоритм ректифікації (1 аркуш А4).
6. Локальний метод пошуку пов'язаних точок (1 аркуш А4).
7. Дослідження локального методу (1 аркуш А4).
8. Пошук диспаратності у вікні (1 аркуш А4).
9. Дослідження алгоритму динамічного програмування (1 аркуш А4).
10. Дослідження напівглобального алгоритма (1 аркуш А4).
11. Достовірність карти глибини (1 аркуш А4).
12. Результати дослідження (1 аркуш А4).
13. Синтез комбінованого алгоритму (1 аркуш А4).
14. Результати дослідження комбінованого методу (1 аркуш А4).
15. Висновки (1 аркуш А4).

КАЛЕНДАРНИЙ ПЛАН

№	Назва етапів роботи	Термін виконання етапів роботи	Примітка
1	Аналітичний огляд літератури	01.09.22–13.09.22	
2	Дослідження відомих методів	14.09.22–27.09.22	
3	Теоретичне обґрунтування	28.09.22–11.10.22	
4	Синтез комбінованого алгоритму	12.10.22–25.10.22	
5	Дослідження комбінованого алгоритму	26.10.22–10.11.22	
6	Графічна частина роботи	11.11.22–25.11.22	
7	Перевірка керівником	26.11.22–02.12.22	
8	Перевірка на академічний плагіат	03.12.22	
9	Перевірка завідувачем кафедри, рецензування	04.12.22–07.12.22	

Дата видачі завдання 01.09.2022 р.

Студент (підпис) Андрій ОСИПЕНКО

Керівник роботи (підпис) Денис ГОРЕЛОВ

РЕФЕРАТ

Пояснювальна записка до атестаційної роботи: 67 сторінок, 27 рисунків, 1 таблиця, 44 джерела.

ДИСПАРАТНІСТЬ, ЕППОЛЯРНІ ЛІНІЇ, КАРТА ГЛИБИНИ, ОБЧИСЛЮВАЛЬНА СКЛАДНІСТЬ, ПАРАЛАКС, РЕКТИФІКАЦІЯ, СТЕРЕОПАРА, ТОЧНІСТЬ

Об'єкт дослідження – процеси оброблення зображень при побудові карт глибини по зображенням стереопари.

Мета дослідження – на прикладі реальних стереопар отримати числові характеристики відомих алгоритмів, які реалізують окремі етапи формування карти глибини, на основі отриманих даних синтезувати комбінований алгоритм який забезпечить високу точність побудови при невеликій обчислювальній складності.

Методи дослідження – аналіз відомих алгоритмів, модельний експеримент, синтез комбінованого алгоритма, статистична обробка експериментальних даних.

В роботі проведено аналітичний огляд методів побудови карти глибини по стереозображенню, досліджено відомі методи побудови карти глибини: методи ректифікації, методи пошуку пов'язаних точок, методи обробки карти глибини, синтезовано комбінований метод побудови карти глибини з підвищеною точністю, синтезований алгоритм реалізовано в середовищі MATLAB, проведено дослідження комбінованого методу з використанням тестових зображень.

ABSTRACT

Explanatory note to the certification work: 67 pages, 27 figures, 1 table, 44 sources.

DISPARITY, EPIPOLAR LINES, DEPTH MAP, COMPUTATIONAL COMPLEXITY, PARALAX, RECTIFICATION, STEREOPORE, ACCURACY

The object of study – the processes of image processing in the construction of depth maps on images of stereo pairs.

The purpose of the study – on the example of real stereo pairs to obtain numerical characteristics of known algorithms that implement the individual stages of formation of the depth map, based on the obtained data to synthesize a combined algorithm that will provide high accuracy with low computational complexity.

Research methods – analysis of known algorithms, model experiment, synthesis of combined algorithm, statistical processing of experimental data.

An analytical review of methods for constructing a depth map is performed stereo image, known methods of depth map construction: methods of rectification, methods of finding related points, methods of depth map processing, combined method of depth map construction with high accuracy, synthesized algorithm implemented in MATLAB environment, research of combined method using test images.

ЗМІСТ

Перелік умовних позначень, символів, одиниць, скорочень і термінів.....	8
Вступ.....	9
1 АНАЛІТИЧНИЙ ОГЛЯД АЛГОРИТМІВ РОЗРАХУНКУ МАПИ ГЛИБИНИ ПО ЗОБРАЖЕННЮ СТЕРЕОПАРИ.....	11
1.1 Постановка задачі.....	11
1.2 Ректифікація стереозображень.....	15
1.3 Виявлення пов'язаних точок.....	17
1.4 Формування мапи глибини.....	22
1.5 Обробка мапи глибини.....	23
1.6 Висновки по розділу 1.....	24
2 ДОСЛІДЖЕННЯ ВІДОМИХ АЛГОРИТМІВ ПОБУДОВИ МАПИ ГЛИБИНИ ПО ЗОБРАЖЕННЮ СТЕРЕОПАРИ.....	25
2.1 Алгоритми ректифікації.....	25
2.1.1 Алгоритм Хартлі.....	25
2.1.2 Метод Малона і Вела.....	27
2.2 Методи пошуку пов'язаних точок.....	28
2.2.1. Локальний метод пошуку.....	28
2.2.2 Використання динамічного програмування для локального алгоритму пошуку.....	33
2.2.3 Застосування динамічного програмування для глобального алгоритму пошуку.....	35
2.2.6 Напівглобальний метод.....	37
2.3 Методи обробки мапи глибини.....	40
2.3.1 Перевірка умови узгодженості диспаратності.....	40
2.3.2 Метод забезпечення субпіксельної точності.....	41
2.3.3 Визначення достовірності.....	41
2.3.4 Інтерполяція.....	43

2.4 Висновки по розділу 2.....	44
3 РОЗРОБКА І ДОСЛІДЖЕННЯ КОМБІНОВАНОГО АЛГОРИТМУ ОБЧИСЛЕННЯ МАПИ ГЛИБИНИ З ПІДВИЩЕНОЮ ТОЧНІСТЮ.....	46
3.1. Обґрунтування алгоритму і його модифікації.....	46
3.2 Комбінований метод побудови мапи глибини.....	47
3.3 Реалізація комбінованого алгоритму в MATLAB.....	50
3.4 Обґрунтування вибору тестових зображень.....	53
3.5 Дослідження комбінованого алгоритму на тестових зображеннях.....	56
3.6 Висновки по розділу 3.....	59
Висновки.....	60
Перелік джерел посилань.....	62
ДОДАТКИ.....	68
Додаток А. Графічний матеріал.....	69
Додаток Б. Відомість кваліфікаційної роботи.....	84

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, ОДИНИЦЬ, СКОРОЧЕНЬ І ТЕРМІНІВ

БПЛА – безпілотний літальний апарат;

BP – Belief propagation – передача повідомлення про сумарний добуток, алгоритм передачі повідомлення для здійснення висновку на графічних моделях;

D – діапазон диспаратності;

GCP – Ground control points – наземні точки прив'язки;

H – висота зображення;

MRF – Markov random field – марківські випадкові поля;

RANSAC – Random Sample Consensus – це ітеративний метод, що використовується для оцінки параметрів математичної моделі для набору спостережуваних даних які містять викиди;

SGM – Semi-Global Matching – напівглобальний метод;

SIFT – Scale-invariant feature transform – масштабнонезалежне перетворення ознак) — алгоритм із області комп'ютерного зору, який виявляє і описує локальні ознаки зображення;

SS – Single Stimulus Methods – методи одиночного стимулювання;

SURF – Speeded up Robust Features – високоефективні детектори / дескриптори інваріантні за масштабом і обертанню особливої точки, для яких стверджується, що вони наближаються або навіть перевершують попередньо запропоновані схеми по відтворюваності, виразності і надійності;

W – ширина зображення;

WTA – Winner-Takes-All – переможець отримує все – застосовується в штучних нейронних мережах при прийнятті рішень і завдань класифікації.

ВСТУП

Технічна довершеність в робототехніці, ігрової індустрії, промисловості і в інших галузях дійшов до такого рівня, що системи комп'ютерного зору, а зокрема стерео зору і фотограметрії, набули широкого поширення і швидко розвиваються. Це відбулося у зв'язку з ростом обчислювальних можливостей і високими завданнями, які стоять перед робототехнічними системами і пристроями, адже за допомогою «зору» робот може отримувати великий обсяг інформації про навколишнє середовище.

Окремим випадком застосування систем технічного зору є навігація рухомих робототехнічних систем, будь то автомобілі Tesla, або марсохід Mars Exploration Rover та багато інших. Також системи технічного зору мають своє застосування в індустрії комп'ютерних ігор, в завданнях виявлення, розпізнавання і розрахунку координат безпілотних літальних апаратів (БПЛА).

Системи комп'ютерного зору, які засновані на застосуванні однієї відеокамери, дозволяють отримувати образи спостережуваних об'єктів, аналізувати кольори та багато інших параметрів. Системи, оснащені 2-ма камерами, тобто системи стереобачення, дозволяють додатково визначати відстань до об'єкта спостереження.

Існує низка проблем стереометоду. Шуми і спотворення на зображення стереопари, відблиски, або протяжні однорідні області в кадрі сильно ускладнюють коректну побудову карти глибини. Тому алгоритм обробки стереозображень повинен мати декілька кроків.

Мета кваліфікаційної роботи – на прикладі реальних зображень стереопар отримати оцінку числових характеристик відомих алгоритмів, що реалізують окремі етапи розрахунку карт глибини, на основі отриманих результатів синтезувати комбінований алгоритм, який зможе забезпечити високу точність карти при незначній обчислювальній складності.

В роботі виконано аналітичний огляд методів розрахунку карти глибини по стереозображенню, досліджено відомі алгоритми розрахунку карти глибини: алгоритми ректифікації, алгоритми пошуку пов'язаних точок, алгоритми обробки карти глибини, синтезовано комбінований алгоритм розрахунку карти глибини з підвищеною точністю, синтезований алгоритм був реалізований в середовищі MATLAB, проведено дослідження синтезованого методу з застосуванням тестових зображень.

В результаті досліджень вдалося отримати помітні результати. Комбінований алгоритм при незначному зростанні складності обчислень забезпечує високу якість карти глибини і достатній рівень деталізації, практично не поступається за цими параметрами відомим високоякісним алгоритмам.

Синтезований алгоритм може бути застосований в роботі різних роботизованих систем технічного зору, які висувають певні вимоги до складності обчислень і точності побудови карти глибини.

1 АНАЛІТИЧНИЙ ОГЛЯД АЛГОРИТМІВ РОЗРАХУНКУ МАПИ ГЛИБИНИ ПО ЗОБРАЖЕННЮ СТЕРЕОПАРИ

1.1 Постановка задачі

В робототехніці, індустрії відеоігор, промисловості велике значення має розрахунок за отриманими фото- або відеозображеннями відстані від об'єктива камери до предметів на зображенні перед камерою.

При використанні роботів у промисловості актуальна задача розрахунку відстані від робочих інструментів робота до тих поверхонь, які вони обробляють. У гірській промисловості частіше за все використовуються роботизовані гірничопрохідницькі автомати, які діють самостійно при участі людини тільки як контролера та спостерігача. Так для правильної роботи робота-екскаватора в шахті необхідна відеосистема, яка дозволяє визначати дистанцію від ковша до пласта породи і від екскаватора до стін шахти задля забезпечення орієнтації екскаватора у тісних умовах шахти.

У комп'ютерних іграх, де замість управління рухом персонажа за допомогою контролерів – клавіатури, миші або джойстика – сам гравець виконує характерні для гри рухи (наприклад, удар ракеткою у грі в теніс). При цьому треба мати дані про дальність об'єкта. Об'єктом може бути сам гравець або ракетка, від якого вимірюється відстань до об'єктиву відеосистеми.

І у випадку робототехніки і у відеоіграх для визначення дистанцій до об'єктів потрібні нескладні з математичної точки зору розрахунки, які дозволяють за мінімальний час визначити відстань з необхідною точністю. Таким методом є розрахунок мапи глибини по дальності.

Мапа глибини – це зображення, на якому кожний піксел несе інформацію не про яскравість або колір об'єкта, а про його відстань до камери. Мапа глибини може бути розрахована за допомогою спеціальних камер глибини, а також може бути обчислена по стереопарі зображень.

Мапу глибини по відстані можна побудувати з використанням ультразвукових перетворювачів або лазерного освітлення, які дають на виході оперативну і точну інформацію про відстань. Якщо не потрібна висока точність або важливий компроміс між ціною і якістю, мапа глибини отримується за допомогою пари стереозображень. Даний метод повторює принцип біноклярного зору людини і дозволяє не викорисовувати численні і дорогі датчики, радари або лідари.

Мапа глибини по відстані формується лише за допомогою двох камер, які, фіксуючи зображення, створюють стереопару. Кожне зображення – це аналог мапини, що людина бачить окремо лівим або правим оком.

Важливими питаннями при побудові мапи глибини по відстані є точність, якість (тобто здатність відрізнити на мапі 2 близько розташованих об'єкта) і час формування мапи.

Система, яка дозволяє будувати мапу глибини по дальності на основі стереозображень (рис. 1.1), включає в себе дві камери, встановлені на одній лінії, що паралельна площині зображення, з відомими дистанціями між камерами – базою B і фокусною відстанню f .

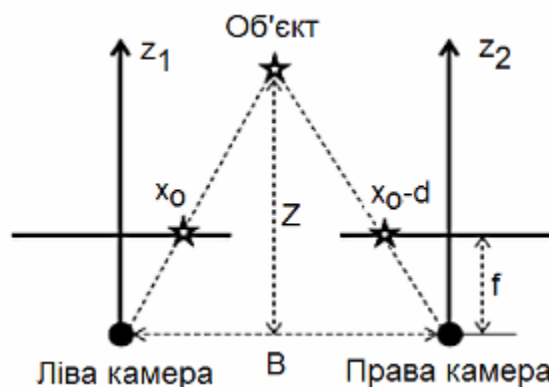


Рисунок 1.1 – Система отримання стереозображень

Також стереосистема може бути побудована на основі 1-єї камери, яка робить знімок з однієї позиції, а потім з другої – переміщеної по лінії, паралельно площині зображення на відому дистанцію.

Однією з фундаментальних проблем стереобаченні є встановлення точної відповідності між правим і лівим зображеннями стереопари. Під відповідністю мається відстань (диспаритет) між пікселями одного і того ж об'єкту на лівому і правому зображеннях. Існують локальні методи по обчисленню мапи диспаритету, засновані на принципах «ковзного вікна» (рис.1.2) [1,2].

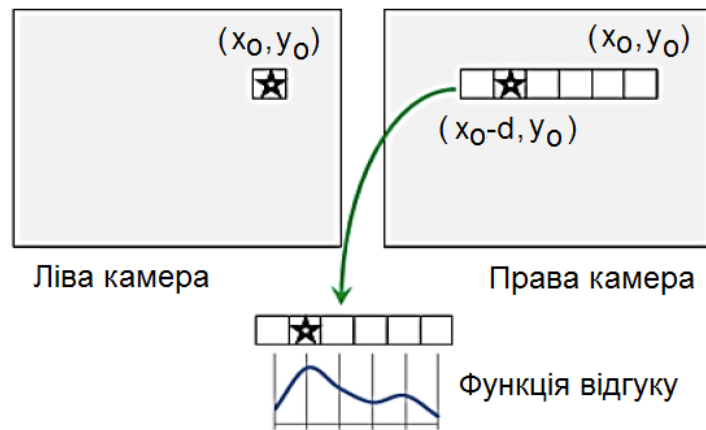


Рисунок 1.2 – Обчислення паралаксного зміщення у стереопарі

Принцип роботи стереосистеми для побудови мапи глибини показано на рис. 1.1. Для кожного пікселя лівого кадра з координатами (x_0, y_0) виконують пошук пікселя у правому кадрі. При цьому вважається, що піксел на правому зображенні має координати

$$(x_0 - d, y_0),$$

де d – величина, що називається невідповідністю або зміщенням.

Пошук відповідного пікселя здійснюється шляхом розрахунку максимуму функції відгуку, в якості неї може виступати, наприклад, взаємна кореляція околиць пікселей [1,2].

Значення глибини є зворотно пропорційним до величини зсуву пікселей. Якщо використовувати позначки з рис. 1.1, то залежність між зміщенням і глибиною можна записати в наступний спосіб:

$$\frac{B-d}{Z-f} = \frac{B}{Z}, \quad \text{тоді } Z = \frac{f \cdot B}{d}, \quad (1.1)$$

де f – фокусна відстань об'єктивів обох камер,

B – база стереосистеми,

d – зміщення видимого положення об'єкта (диспаратність).

Проведення зйомок в умовах недостатнього освітлення, при наявності лише штучних джерел, в яких доводиться працювати обладнанню в реальному житті, призводить до появи шумів на зображеннях стереопари, зроблених камерами. При обробці зображень стереопари потрібно виконувати кореляційний аналіз з врахуванням необхідності застосування кореляційної функції, яка дозволяє побудувати мапу глибини по відстані в незалежності від освітленості і шуму на зображенні. Так нормована кореляційна функція Пірсона дає більш точну мапу глибини по відстані в порівнянні з кореляційною функцією суми квадратів різниць або кореляційною функцією суми модуля різниці, однак в той же час вимагає і більшого часу обчислення.

Для роботи в реальному часі потрібне постійне оновлення мапи глибини, це вимагає пошуку компромісу між якістю мапи і часом її розрахунку.

На рис. 1.1 пошук на іншому зображенні стереопари відповідного фрагмента проводиться за допомогою вікна сканування. Зменшення часу пошуку відповідної області досягається шляхом збільшення розміру вікна. При цьому підбирається розмір вікна сканування, за якого досягається прийнята якість мапи глибини.

Якщо камери, які дають стереопару, встановлені не точно на одній лінії і зміщені по вертикалі одна відносно іншої, перед обробкою обидва зображення мають пройти ректифікацію. Тобто необхідно вирівняти фотографії першу відносно другої так, щоби вибравши фрагмент на одному зображенні, пошук його аналога проводився в тому ж рядку другого знімку. Якщо зображення не пройшли ректифікацію, потрібний фрагмент, пошук

якого виконується вікном сканування, буде зміщено відносно першого зображення по вертикалі. У цьому разі пошук фрагмента треба буде проводити по всьому кадру. Це приведе до збільшення часу розрахунку мапи глибини. Однак сам процес ректифікації вимагає часових витрат і часто потребує наявності людини для контролю процесу.

Тому стереосистема налаштовується таким чином, щоби пошук фрагментів виконувався в однім рядку або з невеликим зміщенням вгору або вниз. У цьому разі досягається оптимальний час обчислення мапи глибини по відстані. При побудови мап глибини проблеми можуть ховатися у складності знаходження відповідних точок у лівому і правому кадрах. Інтенсивність і колір зображення залежать від ракурсу зйомки.

Ретельне налаштування і калібрування камер не гарантує однакову інтенсивність зображення на знімках. Камера генерує цифровий шум, що впливає на отримані кадри. Побудова мапи глибини є доступним інструментом для визначення відстані в незалежності від умов функціонування. Змінюючи ряд параметрів при розрахунку мапи глибини, вдається знайти компроміс при задовільненні вимог, що пред'являються для техніки – якості мапи та швидкості її обчислення.

Отже, у загальному вигляді задачу обчислення мапи глибини можна розділити на такі етапи:

- ректифікація пари стереозображень;
- пошук пов'язаних точок і зіставлення двох зображень;
- обчислення мапи глибини;
- постобробка мапи глибини.

1.2 Ректифікація стереозображень

Якщо припустити, що обидві камери жорстко зв'язані між собою так, що їх оптичні вісі паралельні і знаходяться одна від іншої на певній фіксованій відстані B , то відновлення карти глибини не викликає труднощів.

Однак на практиці обидві камери, за допомогою яких отримують зображення стереопари, дуже складно виставити ідеальним чином. Дві камери треба розташовувати так, щоб в їхнє поле зору потрапляли одні і ті ж самі елементи сцени, і оптичні вісі камер сходилися одна до другої в області розташування предметів сцени [3].

Після отримання пари стереознімків вони мають бути перераховані в єдину прямокутну систему координат з приведенням зображень до епіпольярної стереопари.

Припустимо, є дві камери із центрами проєкцій в точках O і O' , площини проєкцій Π і Π' даних камер і певна точка P в просторі (рис. 1.3) [4].

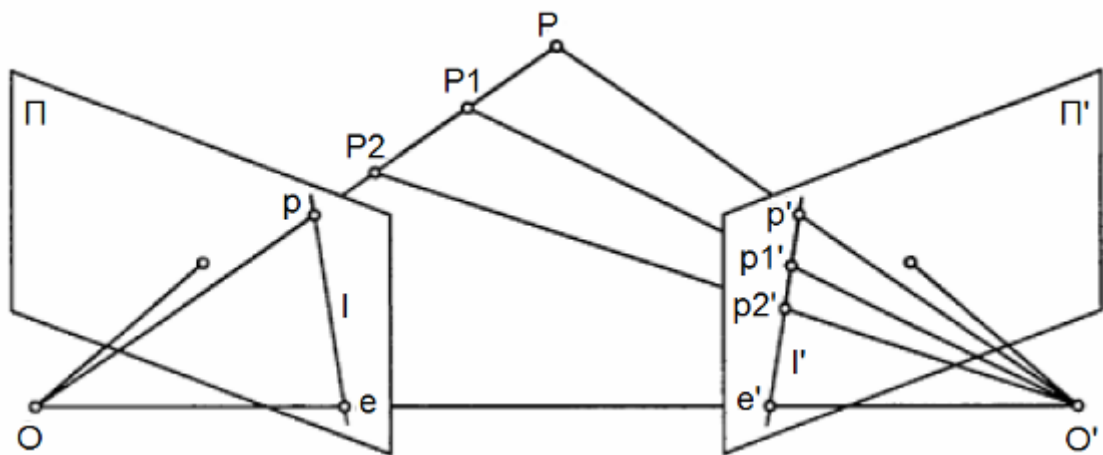


Рисунок 1.3 – Модель пари стереозображень

Називатимемо точки e і e' , у яких лінія OO' перетинає площини Π і Π' , епіполюсами, а прямі l і l' , в яких площина $OO'P$ перетинає площини Π і Π' , епіпольярними лініями для точки P .

Для вектора $t = (t_x, t_y, t_z)$ позначимо кососиметричну матрицю:

$$\mathbf{t}_x = \begin{pmatrix} 0 & -t_z & t_y \\ t_z & 0 & -t_x \\ -t_y & t_x & 0 \end{pmatrix}. \quad (1.2)$$

Матриця t_x має ту властивість, що $t \times l = t_x l$ – є векторним добутком для довільного вектора l [4].

Називатимемо відповідними пару точок на двох зображеннях, які є проєкціями однієї і тієї ж самої точки сцени.

Фундаментальна матриця – це сингулярна матриця розміром 3×3 , що пов'язує відповідні точки. Для відповідних точок x , x' і фундаментальної матриці F виконується умова $(x')Fx = 0$.

Лівий і правий епіполюси є правим і лівим нульовими підпросторами матриці F відповідно для двох кадрів [4].

Об'єкт, що зображений на епіполярній лінії у лівому кадрі, у правому кадрі повинен відображатися також на епіполярній лінії, якщо він в принципі присутній на правому кадрі. В процесі реконструкції стереозображення для ефективного відновлення використовують перетворення, які приводять стереопару до епіполярної. Це перетворення має назву ректифікації і модифікує зображення стереопари так, що для певної системи координат епіполярні лінії стають паралельними до одної з її вісей, і диспаратність виявляється лише вздовж даної осі [5, 6].

Розташування епіполярних ліній на стереозображеннях дає можливість ефективно ототожнювати відповідні об'єкти та елементи сцени на зображеннях стереопари [7].

1.3 Виявлення пов'язаних точок

Задача виявлення відповідності між парами точок на двох стереозображеннях стереопари полягає в ідентифікації характерних фрагментів, які є проєкціями одного і того ж самого фрагменту тривимірного простору. При певних положеннях камери точка поверхні об'єкту може вийти за межі зображення, але якщо її видно на обох зображеннях, то ці обидві точки мають лежати на відповідних епіполярних лініях.

Існує багато підходів до ототожнення точок. В одних методах пропонують аналізувати зображення окремо, виявляючи на них характерні особливості. Такими особливостями можуть бути або об'єкти, які підлягають ідентифікації, або характерні області напівтонового зображення, які ми можемо упевнено впізнати. Зручною для цієї задачі особливістю є виявлення краю. Можна виявити також «кути», де графік функції яскравості має ненульову гаусівську кривизну [3].

В інших методах стереозображення аналізуються спільно, і система намагається кожній точці 1-го зображення поставити у відповідність точку 2-го зображення. Законно очікувати, що між зображеннями яскравість ділянки не змінюється або змінюється незначно в невеликій околиці кожної точки. Можливим є випадок, коли в обох точках рівні яскравості співпадають. Однак реально таких точок існує декілька на відповідних епіполярних лініях [7].

Зазвичай розрізняють два досить широких класи методів для вирішення питання відповідності: локальні і глобальні методи [8]. У підсумку ці два підходи поєднують більшість існуючих рішень питання відповідності [7].

Локальні методи обчислюють диспаратність кожного пікселя окремо. Для цього застосовуються вікна фіксованого або адаптованого розміру, які виділяють блок в опорному (базовому) зображенні. Шляхом обчислення кореляції шукають відповідний блок на другому зображенні [7].

Як правило, в ролі базового опорного зображення виступає кадр з лівої камери, а в ролі другого – з правої. Рахунки відповідності обчислюються як схожість між двома блоками, один з яких знаходиться в базовому кадрі, а другий – в парному. Форма цих блоків залежить від конкретного методу. Стандартні підходи застосовують прямокутне вікно з фіксованим розміром, що визначається експериментально.

Локальні методи відрізняються один від одного використаною метрикою відповідності і методом обчислення рахунків, а саме – розміром околиці сусідніх пікселей, які мають свій внесок.

Навколо довільного пікселя базового кадру будується вікно. На парному зображенні є ряд пікселей, підозрілих на відповідність до заданого пікселя. Навколо кожного такого пікселя будується таке ж саме вікно [7]. Можливі позиції пікселей в парному зображенні задаються мінімально допустимою дистанцією між камерами і об'єктом. Ця дистанція визначає максимальне значення диспаратності. Для кожної з можливих пар вікон проводиться обчислення рахунку відповідності.

Найпоширенішими метриками для обчислення рахунків є квадрат різниці інтенсивностей і абсолютна відмінність інтенсивностей:

$$\begin{aligned}\varepsilon_{SSD}(p, q) &= (L_b(p) - L_m(q))^2, \\ \varepsilon_{SAD}(p, q) &= |L_b(p) - L_m(q)|,\end{aligned}\tag{1.3}$$

де p – піксел базового зображення,

$\mathbf{q} = e(\mathbf{p}, d)$ – піксел парного зображення,

d – величина диспаратності,

$L_b(p)$ і $L_m(q)$ – яскравості пікселей p і q відповідно.

Функція $\mathbf{q} = e(\mathbf{p}, d)$ описує епіполяру лінію на парному зображенні для пікселя p базового зображення з диспаратністю d . Для ректифікованих зображень

$$e(\mathbf{p}, d) = [px - d, py]^T.\tag{1.4}$$

При застосуванні зазначених метрик обчислення рахунків досить ефективне, однак при розрахунках неявно робиться припущення, що для заданого пікселя опорного кадру відповідний піксел парного кадру має приблизно таку ж інтенсивність. Якщо в обох камерах була виставлена різна експозиція або на зображенні, зробленому однією з камер, з'явився відблиск, то описані метрики можуть дати неправильний результат. Крім того, дані методи досить схильні до впливу шумів і випадкових викидів в даних [7].

Дані проблеми в певним чином можна вирішити із застосуванням нормалізованої взаємно-кореляційної функції.

$$\varepsilon_{NCC}(p, q) = \frac{[L_{bp} - \bar{L}_b] \times [L_{mq} - \bar{L}_m]}{\sqrt{\sum_{p \in P} [L_{bp} - \bar{L}_b]^2} \sqrt{\sum_{q \in P} [L_{mq} - \bar{L}_m]^2}}. \quad (1.5)$$

де $\bar{L}$ – середня величина інтенсивності,

P – множина пікселей зображення.

Процес нормалізації зменшує вплив зміни інтенсивності між двома зображеннями шляхом обчислення середнього значення. Ділення на середньоквадратичне відхилення веде значення кореляції до діапазону $[-1, 1]$.

Більшість алгоритмів не враховують дані про колір зображення і керуються лише інтенсивністю L сигналу яскравості, яка має значення в діапазоні від 0 до 255 для 8-бітного сигналу. Додавання проводиться навколо пікселя, який звичайно розташований у центрі вікна. Пошук виконується уздовж епіпольярних ліній [9].

Пара вікон з найбільшою або найменшою, в залежності від алгоритма розрахунків, вартістю і визначають піксел відповідний даному. Але, якщо у вікні порушується умова на неперервність глибини, то його частина впливатиме на результуючу вартість не визначеним чином [10].

Після того, як для кожного пікселя p в базовому зображенні і для кожної з можливих диспаратностей d , від 0 до максимальної, обчислена вартість відповідності, багато локальних алгоритмів переходять до наступного кроку – підсумовування вартостей. Стандартний підхід передбачає у додаванні або усередненні вартостей відповідності у деякій області. Усереднення може виконуватися за допомогою згортки з якимось ядром, найчастіше всього – гаусовим [7].

Вибір форми вікна та його розмірів є важливим питанням. З одного боку, методи кореляції припускають, що функція глибини пікселей всередині

вікна не має розривів. Це означає, що збільшення розміру вікна веде до порушення даної умови і, як наслідок, до невірної роботи методу. З іншого боку, зменшення розміру вікна веде до збільшення впливу шумів, що веде до зменшення вірних збігів.

Головною причиною популярності локальних алгоритмів є їх відносна швидкодія, але вона досягається ціною точності побудови мапи глибини [9].

Метою глобальних методів є пошук найкращої мапи диспаратності відразу для всього зображення. Цю задачу, як правило, вирішують шляхом мінімізації функціонала E . Етап підсумовування рахунків при цьому частіше за все пропускається. Основна задача полягає в знаходженні найкращої конфігурації f , що мінімізуватиме функціонал E :

$$E(f) = E_{data}(f) + \lambda \cdot E_{smooth}(f), \quad (1.6)$$

де $E_{data}(f)$ – вказує, наскільки добре дана конфігурація узгоджується з парою вхідних зображень,

$E_{smooth}(f)$ – відповідає за штраф, який накладається на дану конфігурацію у випадку, коли вона порушує умову неперервності диспаратності,

λ – коефіцієнт, який відповідає за баланс між складовими.

Для обчислення першого доданка часто застосовується попікселне обчислення рахунків за допомогою метрик із локальних методів:

$$E_{data}(f) = \sum C(p, f(p)). \quad (1.7)$$

Для другого доданка частіше за все враховуються сусідні піксели, до яких застосовується певна монотонно-спадна функція, яка знижує рахунки для областей, де інтенсивність міняється значно від пікселу до пікселу.

Коефіцієнт λ визначає ступінь впливу другого доданку на результат мінімізації.

Є цілий ряд методів, які дозволяють знайти мінімум функціоналу E . До них відносяться методи розрізу графа, алгоритми розширення довіри і алгоритми динамічного програмування.

Глобальні методи дають відмінні результати у знаходженні мапи диспаратності. Кількість помилок, які допускаються ними, відносно невелика. Але, всі глобальні методи вимагають значних обчислювальних витрат і часу розрахунку, і жоден з них не зможе бути реалізований у реальному часі [7, 9].

1.4 Формування мапи глибини

Для знайдених пар зв'язаних між собою точок на 2-х зображеннях обчислюється значення диспаратності. Для даного пікселя p базового зображення шуканою диспаратністю буде таке значення d , при якому досягається мінімальне значення рахунку для пари p і d . Цей підхід має назву «переможець отримує все» (Winner-Takes-All). Ці результати формується мапа диспаратності або мапа зсувів, приклад якої наведено на рис. 1.11.

Мапа зсувів застосовується для отримання мапи глибин.

Проблема, яка виникає при даному підході полягає в тому, що для кожної точки обирається унікальна відповідність у парному зображенні. Насправді, одній точці можуть відповідати одразу декілька. Нажаль, така проблема виникає не тільки у локальних методів, але й у переважної більшості алгоритмів, метою яких є побудова мапи диспаратності [7].

Крім того, по причині зворотної залежності глибини і зміщення, роздільна здатність систем стереобачення, які працюють на основі даного методу, є кращою на близьких дистанціях і гіршою на великих.

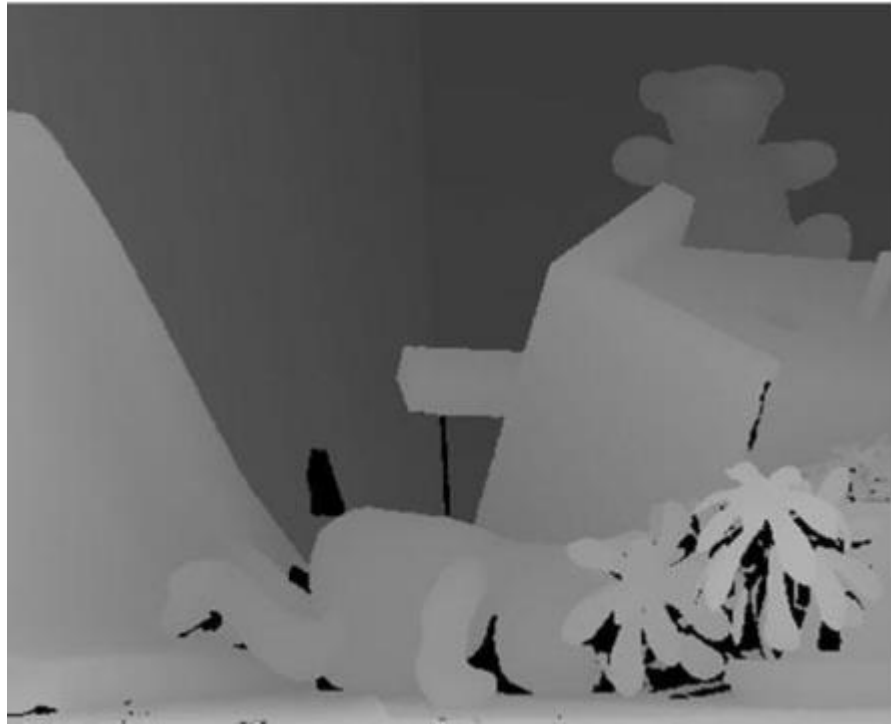


Рисунок 1.4 – Мапа диспаратності

1.5 Обробка мапи глибини

Одним із недоліків більшості алгоритмів побудови мапи глибини є те, що значення диспаратності достовірно виявляється тільки в місцях швидкої зміни яскравості [7]. Застосування методів стереобачення, особливо ретельно оптимізованих і призначених для обробки в реальному масштабі часу на вбудованих процесорах призводить до виникнення помилок при роботі на складних сценах [11]. Це актуально, наприклад, в задачах побудови мапи глибини для малодетальних об'єктів, коли протяжні області на зображенні об'єкта мають приблизно рівні значення яскравості для усіх приналежних до них пікселів і областей з майже нескінченною дальністю [12].

Отже, методи мінімізації помилок, засновані на інтенсивності, можуть враховувати вкрай малі відмінності і визначати довільні значення дальностей для таких областей [11]. Мапа глибини виходить суттєво зашумленою і розрідженою. Така ситуація характерна для обробки сцен, в яких присутня,

наприклад, тінь від об'єкта, рівна стіна, предмет, який має рівномірне забарвлення і малу кількість характерних деталей.

Для оброблювання розріджених і зашумлених мап глибини використовують процедури фільтрації і інтерполяції. При цьому відбувається поширення значення дальностей з ділянок з високою ймовірністю на ділянки з низькою ймовірністю. Останні досягнення в методах технічного зору забезпечують подібну постобробку в умовах реального часу [11].

1.6 Висновки по розділу 1

Метою даної кваліфікаційної роботи є синтез методу побудови мапи глибини по зображеннях стереопари, який забезпечить високу точність розрахунку мапи при компромісі обчислювальної складності.

Для досягнення заявленої мети слід вирішити такі задачі:

- провести аналітичний огляд літератури;
- провести дослідження відомих алгоритмів побудови мапи глибини;
- синтезувати алгоритм побудови мапи глибини, який врахує вимоги до швидкодії і точності отриманої мапи глибини;
- реалізувати алгоритм у програмному середовищі MATLAB;
- виконати експериментальне дослідження синтезованого алгоритму;
- провести аналіз ефективності синтезованого методу.

Область застосувань методів побудови мапи глибини дуже широка – від виявлення та визначення координат дронів [13-18] до забезпечення технічного зору роботів [19,20].

Дана кваліфікаційна робота виконується на кафедрі Медіаінженерії та інформаційних радіоелектронних систем ХНУРЕ. На кафедрі ведуться наукові дослідження, які стосуються технічного зору роботів [19,20], а також виявлення і визначення координат БПЛА за результатами відео- [13-18] і акустичного [21-24] спостереження. Дослідження в даній атестаційній роботі відповідають науковим дослідженням кафедри МІРЕС.

2 ДОСЛІДЖЕННЯ ВІДОМИХ АЛГОРИТМІВ ПОБУДОВИ МАПИ ГЛИБИНИ ПО ЗОБРАЖЕННЮ СТЕРЕОПАРИ

2.1 Алгоритми ректифікації

Парну точку необхідно шукати на епіполярній лінії. Значить, для спрощення пошуку, зображення вирівнюємо так, щоб усі епіполярні лінії були паралельними сторонам кадру (зазвичай горизонтальним). Більш того, зображення вирівнюємо так, щоб у точці із координатами (x_0, y_0) відповідна їй епіполярна лінія описувалася рівнянням $x = x_0$. Тоді для кожної точки відповідну їй парну точку треба шукати на тій же сходинці у кадрі з другої камери. Подібний процес вирівнювання зображень називається ректифікацією (англ. – rectification).

2.1.1 Алгоритм Хартлі

Одним із найбільш популярних алгоритмів ректифікації є метод, запропонований Хартлі [25].

1. Визначаємо відповідності між точками зображень. Можуть застосовуватися будь-які відомі методи виділення характерних точок і будь-які метрики.

2. Обчислюємо фундаментальну матрицю F на основі знайдених відповідностей.

3. Обираємо проєктивне перетворення $H' = G$, що проєктує епіполіус p' на нескінченність (в точку з координатами $(1, 0, 0)$) для епіполіуса $(f, 0, 1)$ і $H' = GRT$ для координат у загальному вигляді, де перетворення R і T – позначають відповідно поворот і паралельний перенос.

Перетворення G записується у вигляді:

$$G = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1/f & 0 & 1 \end{pmatrix}. \quad (2.1)$$

4. Шукаємо відповідне перетворення H для першого кадру. H відповідає H' у випадку

$$Hl = H'l', \quad (2.2)$$

де l та l' – пара відповідних епіполярних ліній,

H^* та H'^* – перетворення площин, матриці даних перетворень є союзними матрицям H і H' .

Перетворення обчислюється шляхом мінімізації суми $\sum_i d(Hu_i, H'u'_i)$ на множині відповідностей точок. При цьому H має вигляд

$$H = (I + H' p' a^t) H' M', \quad (2.3)$$

де a – певний вектор, $F = \mathbf{p} \times M$.

Обирається $H = A$, де

$$A = \begin{pmatrix} a & b & c \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad (2.4)$$

це веде до мінімізації суми $\sum_i (a\hat{u}_i + b\hat{v}_i + c + \hat{u}_i)^2$.

У випадку, коли епіполюс знаходиться всередині кадру, при перетвореннях H і H' він буде розбиватися на дві не зв'язані області (вони визначаються координатою останньої точки у проєктивних координатах), які йдуть у нескінченність, з них обираються ті, що містять відповідні точки.

Області визначаються з такої умови. Припустимо P, P' – матриці проєкцій на зображення, $\{\mathbf{x}_i\}$ – набір точок у 3D просторі, $u_i \Leftrightarrow u'_i$ – відповідні точки у кадрах, що є проєкціями $\mathbf{x}_i$. Тоді якщо $u_i = (u_{i1}, u_{i2}, u_{i3})$ і $u'_i = (u'_{i1}, u'_{i2}, u'_{i3})$, то при усіх i знак $u_{i3} u'_{i3}$ незмінний [25].

Результат роботи методу показано на рис 2.2 у вигляді стереоаналіфа. Оригінальні зображення показані на рис 2.1.



Рисунок 2.1 – Стереозображення Yellowstone



Рисунок 2.2 – Неректифіковане стереозображення і результат його ректифікації у вигляді стереоаналіфа

Чітко помітно розбіжність між зображеннями як в орієнтації, так і в абсолютному положенні. Після ректифікації зображення були перетворені так, що відповідні точки помістилися вздовж одних і тих же рядків.

2.1.2 Метод Малона і Вела

В статті [26] запропоновано інший спосіб знайдення відповідного перетворення H' . Вважається, що $H^{-T} F' H = F$, де $F' = (1, 0, 0)_x$. Отже, є така система:

$$\begin{bmatrix} (h_{21}h'_{31}-h_{31}h'_{21})h'_{31}-h'_{21} \\ (h_{21}h'_{32}-h_{31}h'_{22})h'_{32}-h'_{22} \\ (h_{21}h'_{33}-h_{31}h'_{23})h'_{33}-h'_{23} \end{bmatrix} = \alpha \begin{bmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} \end{bmatrix}. \quad (2.5)$$

Рішення системи (2.5) знаходять методом найменших квадратів. При даних перетвореннях зображень епіполярні лінії стають паралельними вісі x . Після того шукають коефіцієнти у верхніх рядках H і H' .

Розглядається якобіан і ненульові сингулярні числа σ_1 та σ_2 :

$$\mathbf{J}(\mathbf{K}, \mathbf{p}) = \begin{pmatrix} \frac{\partial \bar{x}}{\partial x} & \frac{\partial \bar{x}}{\partial y} \\ \frac{\partial \bar{y}}{\partial x} & \frac{\partial \bar{y}}{\partial y} \end{pmatrix}. \quad (2.6)$$

Якщо $\sigma_1 > 1$, в результаті перетворення будуть створюватися додаткові піксели, якщо $\sigma_1 < 1$ – то відбудеться стискання кадру. Мінімізується величина

$$f(\hat{a}_{11}, \hat{a}_{12}) = \sum_{i=1}^n \left[(\sigma_1(\mathbf{J}(\mathbf{K}, \mathbf{p}_i)) - 1)^2 + (\sigma_2(\mathbf{J}(\mathbf{K}, \mathbf{p}_i)) - 1)^2 \right], \quad (2.7)$$

де $\mathbf{p}_i$ – набір точок в кадрі.

2.2 Методи пошуку пов'язаних точок

2.2.1. Локальний метод пошуку

З метою описання схожості двох пікселей і виявлення сполучених точок застосовуються рахунки відповідності. Для їх розрахунку визначається функція $\varepsilon(p, q)$. Застосовуючи значення інтенсивності, рахунки

відповідності можна розрахувати за допомогою абсолютної різниці яскравостей пари пікселей $\varepsilon_{SAD}(p, q)$:

$$\varepsilon_{SAD}(p, q) = |L_b(p) - L_m(q)|, \quad (2.8)$$

де p – піксел базового зображення,

q – піксел парного зображення.

Для схожих пікселей $\varepsilon_{SAD}(p, q)$ значення буде малим, в той час як для несхожих пікселей воно буде великим [27-29].

За умови того, що кадри відкалібровані і ректіфіковані, пошук сполученого пікселя обмежений одним рядком. Для знаходження відповідності піксел з рядка лівого кадру порівнюється з кожним пікселем відповідного рядка правого кадру.

Діапазон пошуку може бути зменшено спрямовуючими обмеженнями [30]: якщо піксел A на лівому кадру розташований лівіше пікселя B цього ж кадру, то відповідний піксел A_0 на правому кадру буде так само лівіше відповідного B_0 . Однак, ця умова виконується не завжди. Наприклад, вона не виконується для об'єктів малого розміру, розташованих близько до камери (рис. 2.3).

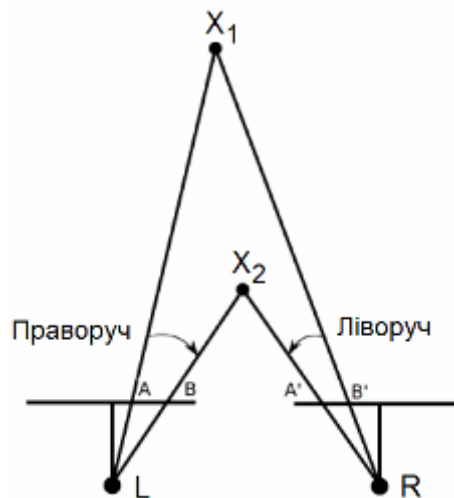


Рисунок 2.3 – Порушення спрямовуючого обмеження для об'єктів малого розміру, розташованих близько до камери

Тут об'єкт X_2 знаходиться правіше X_1 , але оскільки він знаходиться близько до камери і є об'єктом точковим, то зазначена умова порушується.

Спрямовуючі обмеження дозволяють виконувати пошук тільки в заданому діапазоні пікселів і встановити межі значень диспаратності $d \in [d_{\min}, d_{\max}]$.

В результаті пошуку сполученого пікселя може бути отримана кореляційна крива, представлена на рис. 2.4, яка показує мінімум рахунків відповідності для заданого пікселя.

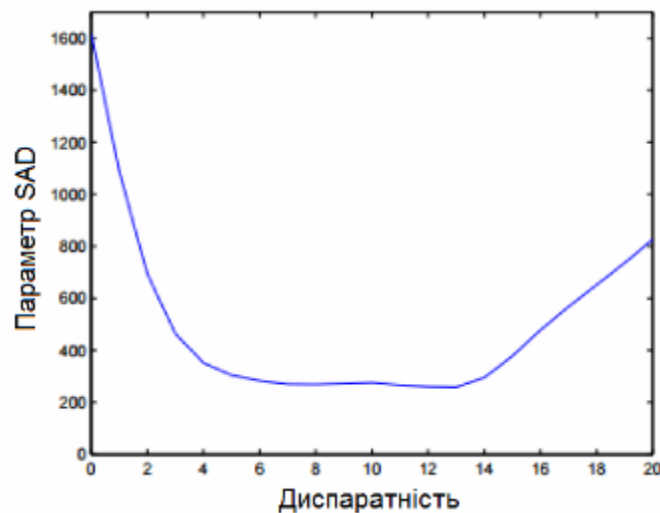


Рисунок 2.4 – Кореляційна крива (мінімум рахунків відповідності)

Потім значення диспаратності обирається методом WTA, без врахування характеру оброблюваної сцени, і в результаті отримується мапа глибини.

Розрахунок рахунків вимагає зробити $O(D)$ кроків для кожного пікселя. Отже, обчислювальна складність локального алгоритму пошуку лінійно зростатиме зі збільшенням кількості пікселів кадру і діапазону диспаратності і виражається як $O(WHD)$, де W – ширина кадру, H – висота кадру, а D – діапазон диспаратності.

Дослідження проведено на відомій стереопарі Tsukuba, що широко використовується для досліджень стереобачення (рис. 2.5). Еталонна мапа глибини і результат роботи методу показані на рис. 2.6.

Не зважаючи на те, що розрахована мапа глибини вийшла дуже детальною, вона в виявилася істотно схильною до шумів. Причиною послужило те, що пошук виконувався попiксельно.



Рисунок 2.5 – Зображення стереопари Tsukuba

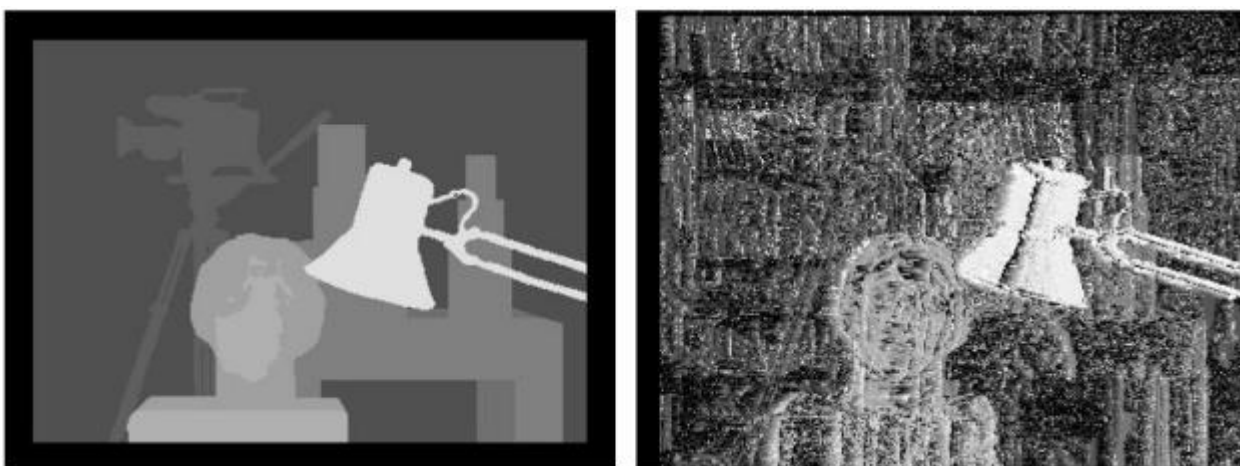


Рисунок 2.6 – Еталонна і експериментальна мапа глибини

Еталонна мапа використовується для визначення точності побудови експериментальної мапи глибини. Методика достатньо проста і зводиться до порівняння значень диспаратності для еталона і отриманої мапи глибини і підрахунку числа пікселів, що не задовольняють певному критерію відповідності. Як правило, в якості критерію використовують абсолютну різницю значень і поріг, який дана різниця не може перевищувати.

Для зменшення шумів пошук можна проводити з врахуванням сусідніх пікселів. Рахунки відповідності можуть бути обчислені за допомогою вікна

розміром $n \times n$, центр якого знаходиться в потрібному для нас пікселі. Тоді наша метрика набуде наступного вигляду:

$$\varepsilon(p, q) = \sum_{\substack{p_N \in N_b \\ q_N \in N_m}} |L_b(p - p_N) - L_m(q - q_N)|, \quad (2.9)$$

де N_b і N_m – вікна відповідно для базового і парного кадрів,

p_N і q_N – піксели всередині цих вікон.

Подібний підхід має на увазі неперервність диспаратності у всьому вікні. На рис. 2.7 показані результати застосування даного методу за різних розмірів вікна.

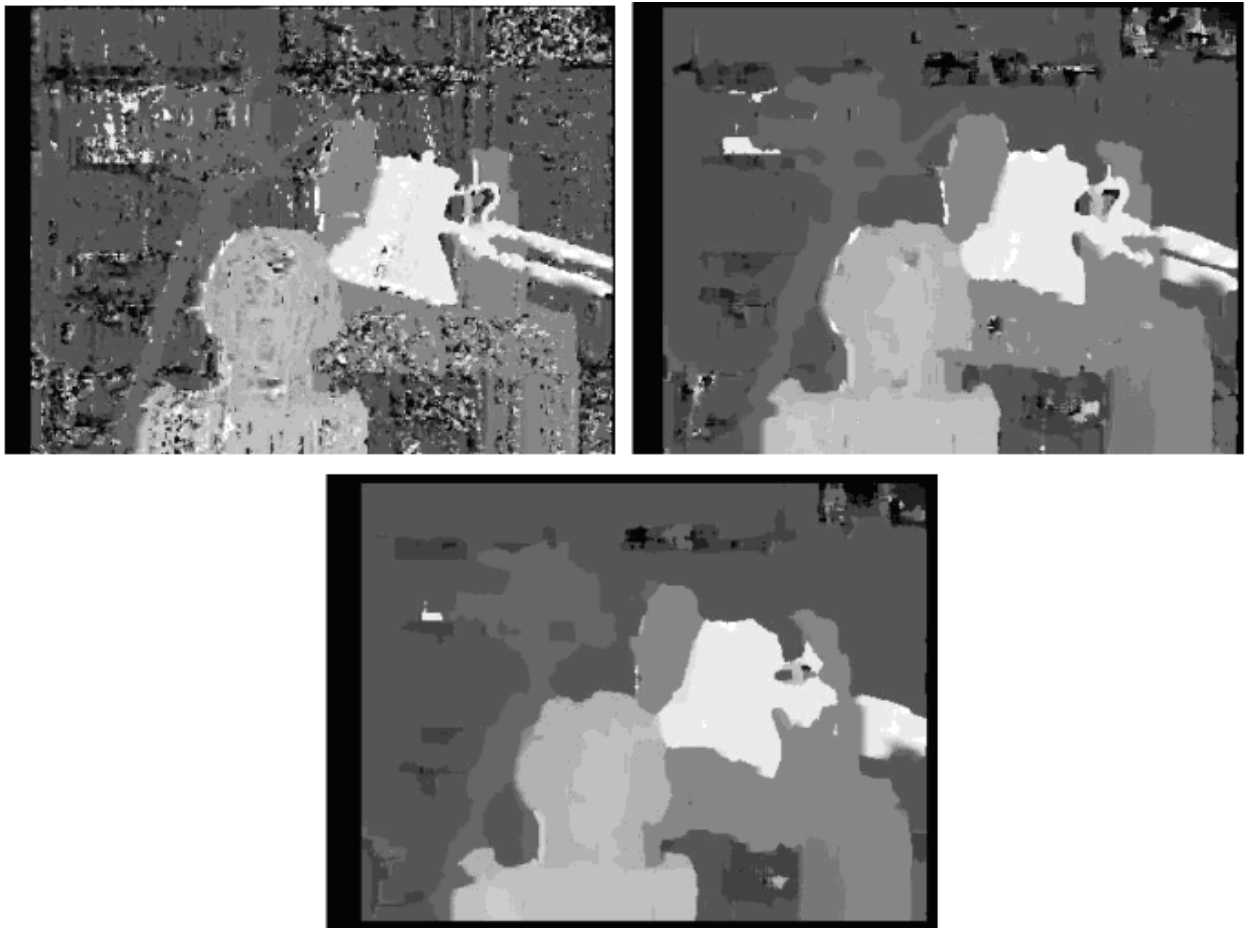


Рисунок 2.5 – Мапа глибини з використанням вікон розмірами 3×3 , 9×9 і 15×15 відповідно

Збільшення розмірів вікна призводить до зниження шумів на мапі глибини, однак різко зменшується кількість деталей на зображенні. Крім

цього, необхідність враховувати сусідні піксели збільшує обчислювальну складність методу.

Важливо зазначити, що локальні методи не справляються з малодетальними та малотекстурованими областями, так як в даному випадку недостатньо інформації для надійного порівняння пікселей в області.

Точність локальних алгоритмів при оптимально підібраних параметрах не вище 73...75%. Якщо є перекриття, результати роботи методу будуть невірними, оскільки він не дає ніяких штрафів на перекриття.

2.2.2 Використання динамічного програмування для локального алгоритму пошуку

У задачах комп'ютерного зору і обробки зображень мають широке використання методи динамічного програмування. У використанні до локальних методів суміщення стереозображень першим етапом алгоритму є відкидання рядків, що відповідають заздалегідь неможливим значенням d . Наприклад, з умови, що піксел у лівому кадрі повинен бути правіше відповідного йому у правому кадрі, і з умови $|d| \leq d_{\max}$.

Використання динамічного програмування зводиться до знайдення оптимального шляху на обчисленій двовимірній матриці. При тому за кожен тип руху призначається деякий штраф. Можливі такі типи руху: по горизонталі, по вертикалі та по діагоналі. Два останні відповідають перекритим областям, що присутні тільки на одному із зображень стереопари.

На рис. 2.8 показана візуалізація роботи описаного методу: горизонтальна ось – ось x , вертикальна – d , чорні піксели відповідають малим, близьким до 0, значенням, стрибки V і D – перекритим областям. Якщо у кадрі заздалегідь відомі контрольні точки, то кількість можливих шляхів різко зменшується за рахунок введення обмеження: шлях повинен проходити таким чином, щоби він не суперечив розстановці контрольних точок.

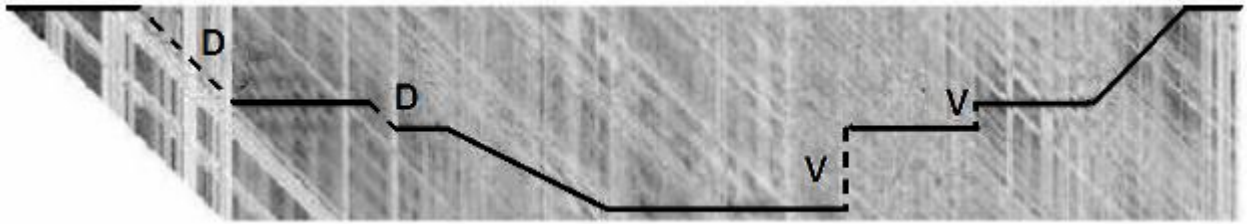


Рисунок 2.6 – Пошук оптимального шляху на обчисленій двовимірній матриці

Під контрольними (або GCP) точками, розуміють положення точок, відповідність яких ми можемо обчислити досить точно. Можливі також багатозначні точки GCP, коли точці одного кадру може відповідати одна точка з невеликого набору варіантів у другому кадрі [31].

Результат роботи алгоритму показано на рис. 2.9.



Рисунок 2.9 – Результат застосування алгоритму динамічного програмування

Точність побудови мапи глибини для даного методу сягає до 85%.

Значно підвищити швидкість роботи дозволяє те, що кожний рядок кадру обробляється незалежно. Певний внесок вносить і можливість паралельних обчислень. Але, в результаті функція глибини піксела може мати вигляд гребінки.

Із використанням динамічного програмування [30], обчислювана складність локального методу знижується порівняно із загальним алгоритмом і не залежить від діапазону диспаратності $O(WH)$.

2.2.3 Застосування динамічного програмування для глобального алгоритму пошуку

В роботі [31] пропонується застосувати ідеї динамічного програмування, щоби мінімізувати глобальний функціонал:

$$E(f) = E_{data}(f) + \lambda E_{smooth}(f),$$

$$E_{data}(f) = \sum_{p \in \mathbf{P}} m_f(p), \quad (2.10)$$

де $m_f(p) = \min \varepsilon_{SAD}(p, q)$,

$\mathbf{P}$ – множина пікселей в лівому кадрі, p відповідає q згідно конфігурації f .

Позначимо $G = (P, N)$ – граф з вершинами-пікселями. Кожен піксел (крім крайніх) з'єднано з чотирма своїми сусідами (отже N – множина сусідніх пікселей). Значить

$$E_{smooth}(f) = \sum_{(p, p_N) \in N} s(d(p), d(p_N)), \quad (2.11)$$

де $s(d_1, d_2)$ – функція гладкості, що ставить у відповідність парі значень диспаратності штраф за значну відмінність в їхніх значеннях.

Якщо ми певним чином оберемо дерево-підграф G (з виділеним коренем $root$) – $G_0 = (P, N_0)$, то метод мінімізації буде виражений рекурентною формулою:

$$E_p(d(par(p))) = \min_{d(p) \in D} (m_f(p) + s(d(p), d(par(p)))) +$$

$$+ \sum_{p_N \in child(p)} E_{p_N}(d(p)). \quad (2.12)$$

При тому мінімум береться по усіх можливих присвоюваннях пікселу-вузлу p диспаратності $d(p)$, D – є множиною можливих розбіжностей.

Якщо $p = \text{root}$ – корінь, то доданок $s(d(p), d(\text{par}(p)))$ вважається рівним нулю. В процесі мінімізації $E_{\text{root}}(d(\text{root}))$ для кожного пікселя p обирається розбіжність $d(p) \in D$, так визначається оптимальна конфігурація f .

В якості G_0 в роботі [31] пропонуються 2 варіанти: MID Tree і MDDT Tree.

MID Tree – є найменшим остовним деревом G , побудованим з того розрахунку, що вага ребра складає $v_{pp_N} = |L(x) - L(p_N)|$.

Оскільки багато ребер матимуть однакові ваги, то отримане дерево не є унікальним. Є інший принцип побудови дерева: для кожної точки обчислюється функція $D(x)$, яка пікселу ставить у відповідність величину, наскільки далеко від границь він міститься у певній однорідній (наприклад, по інтенсивності) області. В цьому випадку вага

$$u_{pp_N} = d_{\max} - \frac{1}{2}(D(p) + D(p_N)), \quad (2.13)$$

$$\text{де } d_{\max} = \max_{p \in P} d(p).$$

Тоді

$$D(X) = \max_{\hat{x} \in B} \text{dist}(x, \hat{x}). \quad (2.14)$$

На рис.2.10 показано результат роботи алгоритму динамічного програмування. Визначена точність даного методу дорівнює 86%.

Можна обирати ваги не v_{pp_N} чи u_{pp_N} , а їх комбінацію

$$c_{pp_N} = kv_{pp_N} + u_{pp_N}, \quad (2.15)$$

$$\text{де } k = \max_{(p, p_N) \in N} u_{pp_N}.$$

Мінімальне остовне дерево, що побудоване з урахуванням даних вагів, називається MDDT [31].



Рисунок 2.10 – Результат застосування методу динамічного програмування

Згідно роботи [32], обчислювальна складність методу дорівнює $O((WH)^3 \lg(WH))$.

2.2.6 Напівглобальний метод

Метод напівглобального зіставлення або Semi-Global Matching було запропоновано Хіршмюллером в 2005 році в роботі [33]. Він використовує ідею попиксельного пошуку пов'язаних точок і апроксимацію глобальної двовимірної функції гладкості при комбінуванні множини одновимірних функцій [33].

Рахунок відповідності обчислюється для пікселя **p** базового кадру на підставі його яскравості L_{bp} і вірогідної відповідності L_{mq} для пікселя **q** парного кадру.

Один із способів попиксельного обчислення рахунків полягає в алгоритмі, запропонованому Берчфілдом і Томасі [34]. Рахунок $C_{BT}(p, d)$ обчислюється як абсолютна мінімальна різниця яскравостей у точках **p** і **q** в діапазоні $1/2$ пікселя в кожному напрямку уздовж епіполлярної лінії.

Альтернативою є обчислення рахунків відповідності, засноване на взаємній інформації MI [35]. Вона визначається із ентропії H двох кадрів і їхньої взаємної ентропії:

$$MI_{L_1, L_2} = H_{L_1} + H_{L_2} - H_{L_1, L_2}, \quad (2.16)$$

Ентропії обчислюються з розподілу інтенсивностей P оброблюваних кадрів:

$$H_L = - \int_0^1 \int_0^1 P_L(i) \log P_L(i) di, \quad (2.17)$$

$$H_{L_1, L_2} = - \int_0^1 \int_0^1 P_{L_1, L_2}(i_1, i_2) \log P_{L_1, L_2}(i_1, i_2) di. \quad (2.18)$$

Для задач стереовідповідності один з кадрів має бути перетворений у відповідності з мапою диспаратності для досягнення умови відповідності другому кадру, щоб пов'язані піксели у кожному кадрі перебували в одній і тій самій позиції, тобто

$$L_1 = L_b, \quad L_2 = f_D(L_m). \quad (2.19)$$

З метою розв'язання цієї задачі було запропоновано ітераційний підхід, в якому перша ітерація починається з генерації довільної мапи диспаратності для розрахунку SMI . Отримані рахунки використовуються далі для зіставлення двох зображень і розрахунку нової мапи диспаратності, яка є основою для наступної ітерації. Для якісної роботи методу достатньо трьох ітерацій, так як навіть помилкова мапа диспаратності дозволяє зробити гарну оцінку розподілу P .

Крім цього, обчислення починається для зображень, зменшених до 1/16 від оригінального розміру. У кожній наступній ітерації мапа глибини

масштабується і використовується у розрахунках. Якщо загальна складність методу $O(WHD)$, то час виконання для половинного розрізнення зменшується в $2^3 = 8$ разів [33]. Для 3-х ітерацій при розрізненні 1/16 від оригінала і наступному його збільшенні час розрахунку зростає незначно:

$$1 + \frac{1}{2^3} + \frac{1}{4^3} + \frac{1}{8^3} + 3\frac{1}{16^3} \approx 1,14.$$

Попіксельне обчислення рахунків в основі своїй неоднозначне, і неправильні відповідності можуть мати меншу вагу, ніж правильні, зважаючи на наявність шумів, рівномірних ділянок, відблисків і тін. У зв'язку з цим вводять додаткове обмеження, що забезпечує гладкість шляхом призначення штрафу для зміни диспаратності у сусідніх пікселях:

$$\begin{aligned} E(d) = \sum_{\mathbf{p}} C(\mathbf{p}, d(\mathbf{p})) + \sum_{p_N \in N_{\mathbf{p}}} P_1 \mathbf{T}[|d(\mathbf{p}) - d(p_N)| = 1] + \\ + \sum_{p_N \in N_{\mathbf{p}}} P_2 \mathbf{T}[|d(\mathbf{p}) - d(p_N)| > 1], \end{aligned} \quad (2.20)$$

де $\mathbf{T}[]$ – оператор, що дорівнює 1, якщо його аргумент є «істина», і 0, якщо «неправда».

Перший член виразу (2.20) є сумою всіх рахунків для диспаратностей d . Другий член у (2.20) накладає штраф P_1 для усіх пікселей q в околиці $N_{\mathbf{p}}$ пікселя $\mathbf{p}$, для яких диспаратність небагато змінюється. Третій член (2.20) накладає більший штраф P_2 на більшу зміну диспаратності. Застосування меншого штрафу для менших змін забезпечує адаптацію методу до похилих і кривих поверхонь. Великий штраф для усіх значних змін забезпечує захист від розривів [36]. Розриви виглядають як зміни інтенсивності.

На рис 2.11 показано результат роботи методу SGM.



Рисунок 2.11 – Результат роботи алгоритму SGM

Отже, розглянутий алгоритм дає набагато кращі результати, ніж локальні алгоритми, і по точності може порівнятися із глобальними алгоритмами (точність досягає 92%). При цьому, алгоритм SGM набагато швидше глобальних алгоритмів.

2.3 Методи обробки мапи глибини

2.3.1 Перевірка умови узгодженості диспаратності

Перевірка умови узгодженості диспаратності [37-40] застосовується для виявлення розбіжностей в мапах глибини для лівого і правого кадрів. Перевірка виконується за рахунок зміни ролей для лівого і правого кадрів в методі. Для обчислених мап лівого і правого кадрів проводиться звірка значень диспаратності для зв'язаних пікселей і, якщо ці значення не збігаються, значення приймається неправильним (необ'єктивним). Дану умову можна записати наступним чином [27]:

$$d(p) = \begin{cases} d_b(p), & |d_b(p) - d_m(q)| \leq 1, \\ d_{inv}, & |d_b(p) - d_m(q)| > 1. \end{cases} \quad (2.21)$$

Описана перевірка особливо ефективно виявляє перекриті ділянки зображення в кадрах [33].

2.3.2 Метод забезпечення субпіксельної точності

Обчислення диспаратності дає дискретні значення глибини. Для усунення ефекту дискретизації, можна здійснювати субпіксельне оцінювання, засноване на квадратичній інтерполяції [70].

Квадратична крива проводиться крізь значення рахунків, сусідніх з мінімальними, і визначається позиція мінімуму функції апроксимації [33]. Обчислене значення мінімуму приймається за значення диспаратності в заданій точці.

2.3.3 Визначення достовірності

Достовірність визначимо як функцію (p, d) . На рис. 2.12 наведено характерний вид кривої кореляції.

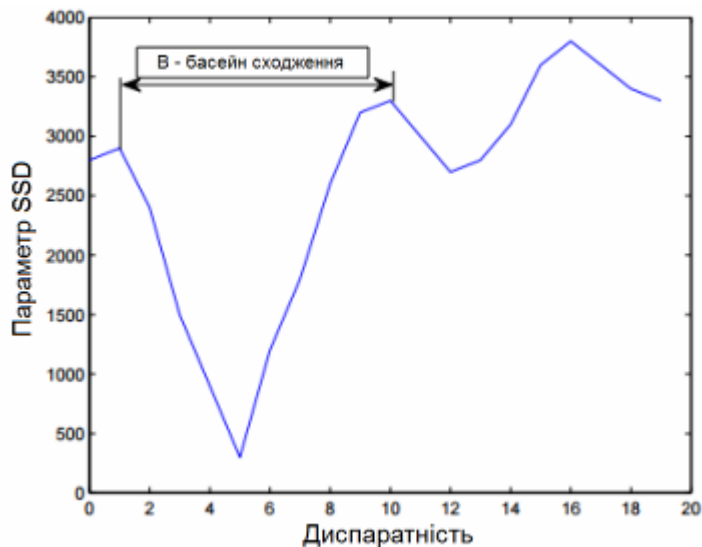


Рисунок 2.12 – Крива кореляції і басейн сходження

Знаючи величину диспаратності d , пропонується розраховувати достовірність у вигляді:

$$C(d) = \frac{B(d)}{d_{\max} - d_{\min}}, \quad (2.22)$$

де $C(d)$ – достовірність поточної диспаратності,

$B(d)$ – «басейн» сходження оцінки величини диспаратності,

$d_{\max} - d_{\min}$ – діапазон величини диспаратності.

Очікується, що на однорідних малодетальних ділянках крива буде мати багато локальних мінімумів із малими значеннями $B(d)$. І оскільки $C(d)$ пропорційне $B(d)$, то слід очікувати малих достовірностей. У разі, якщо локальний мінімум є одночасно і глобальним мінімумом, достовірність буде максимальною.

Значення «басейну» сходження $B(d)$ визначається аналогічно методом градієнтного спуску. Приймаючи за оцінку диспаратності d , можна очікувати, що градієнт праворуч від d буде позитивним, і ліворуч – від'ємним. Алгоритм покроково виконує рух в обох напрямках до тих пір, поки знак градієнта не зміниться, відзначаючи локальний максимум ліворуч і праворуч від d . Діапазон між даними локальними максимумами і є $B(d)$. Після розрахунку усіх достовірностей задається порог T , який відсіває значення диспаратності із меншими достовірностями.

На рис. 2.13 показано результат застосування алгоритму визначення достовірності до мапи глибини при $T = 2/15$.

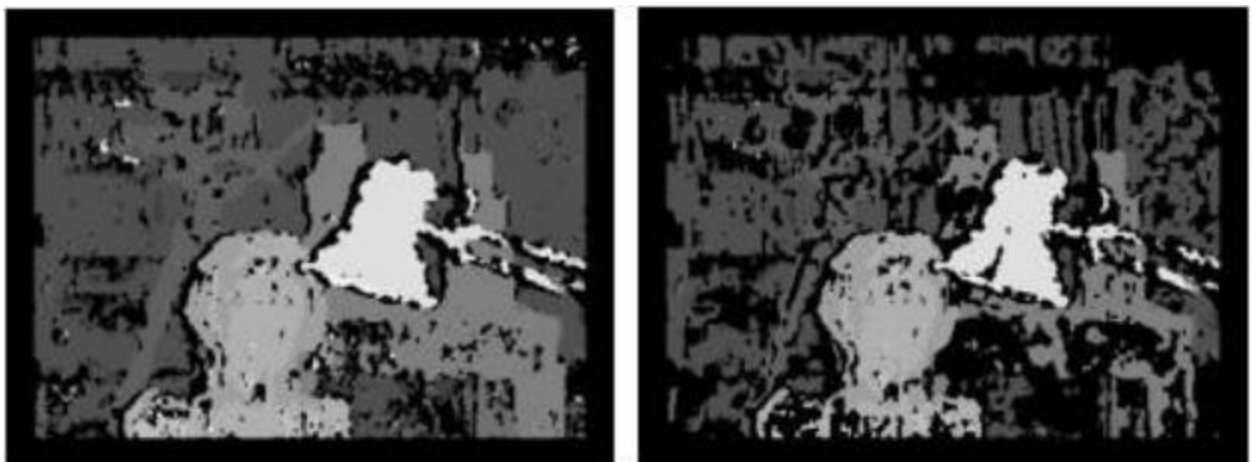


Рисунок 2.13 – Мапа глибини до і після обробки алгоритмом визначення достовірності

2.3.4 Інтерполяція

Актуальною є проблема побудови мапи глибини для однорідних малодетальних об'єктів, областей з майже нескінченною дальністю і областей, які містять відблиски [42, 43]. Мапа глибини виявляється в значно зашумленою і розрідженою. У таких випадках може знадобитися інтерполяція значень мапи глибини з метою покращення деталізації.

Свою ефективність показало використання алгоритмів, що здійснюють зіставлення меж мапи глибини за межами вихідного зображення і поширення величину дальностей з ділянок з великою достовірністю на ділянки з малою достовірністю [11]. Таким алгоритмом є метод швидкого глобального згладжування, заснований на зваженому алгоритмі найменших квадратів [11].

Алгоритм основний на глобальних методах і здійснює мінімізацію функціоналу зваженого метода найменших квадратів. В загальному випадку, в задачі згладжування зі збереженням границь функціонал має такий вигляд:

$$J(u) = \sum_p \left((u_p - f_p)^2 + \lambda \sum_{q \in N(p)} w_{p,q}(g) (u_p - f_q)^2 \right), \quad (2.23)$$

де f – вхідний кадр,

g – керівний кадр,

u – очікуваний результат,

f_p і g_p – значення яскравості пікселя p .

Керівний кадр g може бути заданий вхідним кадром f (для задач фільтрації зображень) або іншим керівним сигналом, узгодженим із f , в залежності від використання.

$S = W \times H$ – розмір кадру,

$N(p)$ околиця пікселя p – набір сусідніх пікселів, як правило 4,

λ – коефіцієнт, що відповідає за баланс між двома складовими, збільшення величини λ призводить до більшого згладжування ефекту u .

Обмеження на згладження підсилюється ваговою функцією $w_{p,q}(g)$, що представляє схожість між двома пікселями p і q .

2.4 Висновки по розділу 2

В результаті проведеного дослідження встановлено, що найбільший вплив на складність обчислень і точність формування мапи глибини має етап пошуку зв'язаних точок. Саме в цій області ведеться більшість розробок, задачею яких є оптимізація тих чи інших характеристик методів 3D реконструкції.

В зведеній табл. 2.1 показані характеристики досліджених алгоритмів.

Таблиця 2.1 – Характеристики для досліджених алгоритмів побудови мапи глибини

Метод	BM	BM-DP	GM-DP	SGM
Обчислювальна складність	$O(WHD)$	$O(WH)$	$O((WH)^3 \cdot \lg(WH))$	$O(WHD)$
Точність	73%	85%	86%	92%

Виконаємо порівняльний аналіз. Дослідження показали, що найкращу точність визначення стереовідповідності забезпечує алгоритм розрізу графа. В той же час він же має найбільшу складність обчислювань.

Метод локального пошуку сильно поступається в точності обчислення 3D моделі, кількість помилок для нього найвища. Незважаючи на те, що своє застосування він отримав завдяки швидкості своєї роботи, ряд методів на даний момент практично не поступається йому, і забезпечують при цьому значно менший відсоток помилок. До таких методів в першу чергу

відносяться алгоритм динамічного програмування (оснований на локальному методі) та напівглобальний алгоритм.

Значний вплив на результат можуть мати алгоритми обробки мапи глибини. Їхнє використання не завжди виправдане, тому що ці алгоритми часто збільшують складність обчислювання усього алгоритму. Але при використанні алгоритмів пошуку з порівняно невисокою складністю обчислень можна досягти значущих результатів у збільшенні якості мапи глибини.

Відомі методи ректифікації у своїй більшості мають приблизно однакову швидкість роботи і дають подібні результати. Застосування більш складних методів не виправдано взагалі, оскільки головна мета ректифікації – спростити обчислення на етапі пошуку пов'язаних точок – втрачає сенс при значній обчислювальній складності самого методу ректифікації.

3 РОЗРОБКА І ДОСЛІДЖЕННЯ КОМБІНОВАНОГО АЛГОРИТМУ ОБЧИСЛЕННЯ МАПИ ГЛИБИНИ З ПІДВИЩЕНОЮ ТОЧНІСТЮ

3.1. Обґрунтування алгоритму і його модифікації

В результаті проведеного у розділі 2 дослідження відомих алгоритмів побудови 3D моделі в даному розділі виконуємо обґрунтування і опис комбінованого рішення з обчислення мапи глибини.

У якості методу ректифікації вхідних кадрів пропонується застосувати метод Хартлі. Метод складається з 4 етапів:

- виявлення особливих точок,
- зіставлення особливих точок на стереопарі,
- розрахунок фундаментальної матриці,
- факторизація фундаментальної матриці.

Виявлення особливих точок доцільно виконувати за допомогою методу SURF [32], розрахунок фундаментальної матриці – за допомогою методу RANSAC [74].

Алгоритм SURF позитивно зарекомендував себе у задачах пошуку об'єктів на зображеннях, 3D-реконструкції, у порівнянні кадрів. У задачах ректифікації корисною є його властивість інваріантності до зміни масштабу і повороту.

Алгоритм RANSAC в свою чергу здатний надати стійку оцінку фундаментальної матриці і відмітити відповідності як викиди чи не викиди. При застосуванні цього метода шукається єдиний набір даних, найкраще відповідний моделі в конкретному сенсі. Популярність цього методу пов'язана з високою його стійкістю до грубих помилок типа збоїв. На сьогодні він визнаний одним із найбільш ефективних в умовах сильного шуму у вхідних кадрах [75].

Для здійснення пошуку зв'язаних точок на підставі дослідження, проведеного у розділі 2.2, було обрано напівглобальний метод, який

заснований на ідеї попиксельного порівняння взаємної інформації та апроксимації глобальної 2D функції гладкості шляхом комбінації множини одновимірних функцій. Так як рахунки відповідності, засновані на використанні метрик для яскравостей зображень чутливі до відмінностей в умовах зйомки і освітлення, відбиттів і відблисків [63], то вибір був зроблений на користь застосування взаємної інформації.

Взаємна інформація дозволяє порівнювати зображення зі складними залежностями відповідних яскравостей, у тому числі і зображень камер різного типу [62], [64]. Переваги даного підходу, в порівнянні з іншими методами, полягають у тому, що, з одного боку, він дозволяє досягти високої якості при обчисленні мапи глибини, сильно випереджаючи локальні методи, а з іншого боку, швидкість його роботи набагато вище, ніж у глобальних методів.

Мапа глибини формується алгоритмом WTA. З метою зменшення кількості помилок для отриманої мапи глибини і суміщення кордонів 3D моделі з межами об'єкту у базовому кадрі пропонується використати інтерполяцію алгоритмом швидкого глобального згладжування. Описаний алгоритм є ефективним інструментом обробки розріджених мап глибини, тому що дозволяє збільшити точність побудови мапи глибини для сцен, які містять однорідні об'єкти, і забезпечує при цьому високоякісну обробку. У той же час, метод має лінійну складність обчислення, що важливо для роботизованих систем, вимогливих до швидкодії.

3.2 Комбінований метод побудови мапи глибини

На основі попередніх досліджень було розроблено алгоритм побудови мапи глибини, показаний на рис. 3.1.

Алгоритм Хартлі, що використовується для ректифікації, у повній мірі проаналізовано в 2.1.1. Застосований для виявлення особливих точок метод SURF добре відомий і досить докладно викладено в багатьох роботах [32].

На обох кадрах виділяються характерні точки і обираються потенційні відповідності. При тому частина точок може не мати відповідності у парному кадрі. Правильно суміщені точки мають задовольняти обмеженням епіполярних ліній. Це означає, що точка повинна лежати на епіполярній лінії, визначеній зв'язаною точкою. За цією умовою проводиться розрахунок фундаментальної матриці за допомогою методу RANSAC. Далі обирається проєктивне перетворення і виконується вирівнювання зображень.

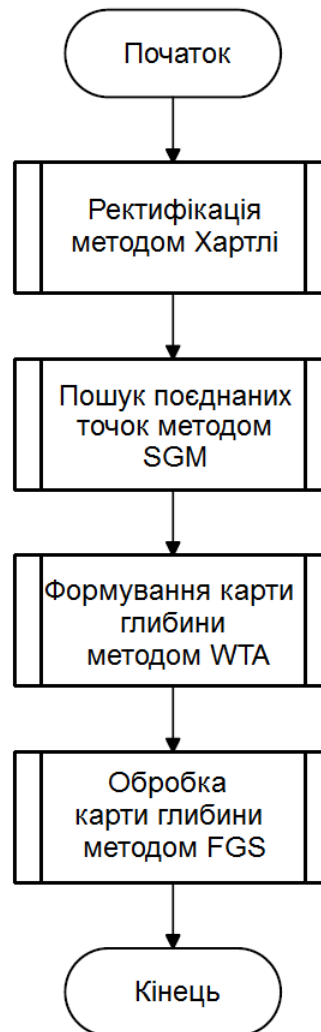


Рисунок 3.1 – Схема комбінованого алгоритму

Наступним етапом побудови 3D моделі є пошук зв'язаних точок за допомогою напівглобального методу, що заснований на попільсельному зіставленні взаємної інформації та апроксимації глобальної 2D функції гладкості шляхом комбінації множини одновимірних функцій [32]. Основний алгоритм повторює дії, викладені в 2.2.6.

WTA є універсальним алгоритмом формування мапи глибини, його суть повністю викладено у розділі 1. Для отриманої мапи глибини використовується метод швидкого глобального згладжування (розділ 2.3.4), який засновано на зваженому алгоритмі найменших квадратів, він виконує неоднорідне згладжування зі збереженням границь об'єктів.

Популярний підхід до роздільного згладжування полягає в послідовному використанні одновимірних рішень вздовж кожного вимірювання зображення [40]. Для заданого 2D зображення одновимірне рішення застосовується ітераційно вздовж рядків і стовпців кадру. Але цього недостатньо для поширення інформації вздовж границь, і як результат – виникають «смугові артефакти», які властиві роздільним методам.

Щоб цього уникнути, 2D згладжування здійснюється методом багаторазового застосування одновимірного згладжування зі змінним коефіцієнтом згладжування для кожної ітерації. Сумарний коефіцієнт розраховується за формулою:

$$\lambda = \sum_{t=1}^T 2\lambda_t = \sum_{t=1}^T (1/2)^{2t} \lambda_t, \quad (3.1)$$

де λ_t – коефіцієнт згладжування для ітерації за номером t .

Тоді, якщо задатися необхідним кінцевим значенням коефіцієнта, можна отримати значення для кожного проходу фільтра:

$$\lambda_t = \frac{3}{2} \frac{4^{T-t}}{4^T - 1}. \quad (3.2)$$

Мінімальне число ітерацій, які дають результати обробки, що не поступаються за якістю глобальним алгоритмам, було встановлено в експерименті. Дане число дорівнює 3.

3.3 Реалізація комбінованого алгоритму в MATLAB

Відповідно до описаного в розділі 3.2 методу, реалізація відповідного алгоритму в MATLAB має з декілька блоків обробки.

Перший блок – отримання вхідних даних і параметрів алгоритму. За допомогою функції `imread` виконується зчитування стереозображень з файлів, розташованих в папці `/data`. Потім здійснюється приведення кадрів до монохромного вигляду за допомогою функції `rgb2gray`. В програмі передбачається можливість масштабування вхідних зображень з використанням функції `imresize`.

Для керування параметрами пошуку мапи глибини задається діапазон значень диспаратності `maxDisparity`, коефіцієнти згладжування $P1$ і $P2$ для алгоритму SGM, коефіцієнт згладжування `lambda` і кількість ітерацій `NumIt` для алгоритму FGS. Функція `size` використовується, для збереження розмірів вхідних зображень, що знадобиться для подальшої обробки.

Другий блок – функція ректифікації. Вона реалізована викликом зовнішньої функції `stereoRectify`, яка приймає на вхід пару некаліброваних кадрів і повертає результат їх ректифікації. В свою чергу в цій функції реалізуються етапи алгоритму Хартлі. Спочатку виконується виявлення особливих точок за допомогою методу SURF [32], реалізованого у вбудованій функції `detectSURFFeatures`. За допомогою функції `selectStrongest` здійснюється відбір характерних точок з найбільш сильними метриками. Функція `matchFeatures` виявляє передбачувані пов'язані точки. Розрахунок фундаментальної матриці виконується за допомогою методу RANSAC. Застосування функції `estimateUncalibratedRectification` дозволяє робити оцінку перетворень, потрібних для ректифікації стереопари, що застосовуються в функції `rectifyStereoImages`.

Третій блок – функція пошуку пов'язаних точок. Вона так само реалізований у зовнішній функції `stereoSGfunc`. На вхід цієї функції

передаються ректифіковані увдри і параметри, які задають максимальну диспаратність і коефіцієнти згладжування.

У тексті даної функції реалізуються етапи напівглобального методу пошуку. Перший етап – ініціалізація початкової мапи глибини у вигляді масиву розмірністю, рівною до розмірності лівого кадру, заповненого нулями. Дане наближення є достатнім, оскільки навіть нульова мапа глибини дозволяє виконувати оцінку розподілу інтенсивностей досить точно.

Потім реалізується цикл з декількох ітерацій алгоритму. На першому етапі знижується розмірність обох зображень і мапи глибини в 8 разів.

Для обчислення рахунків відповідності викликається зовнішня функція `stereoSGcost`, на вхід якої надходять масштабовані масиви лівого і правого кадрів і мапи глибини, а також відповідно зменшеного значення максимальної диспаратності.

Дана функція, в свою чергу, реалізує кроки обчислення рахунків відповідності. Спочатку виконується ініціалізація нулями масиву розміром 256×256 , в який буде записаний взаємний розподіл інтенсивностей для пари зображень, і масиву розміром $W \times H \times D$, в якому зберігатимуться рахунки відповідності кожного пікселя для кожного значення диспаратності. Потім виконується розрахунок взаємного розподілу шляхом обчислення усіх можливих пар відповідностей і поділення результату на загальну кількість відповідностей.

Потім виконується підсумовування масиву розподілу по рядкам і по стовпцям з метою отримання власного розподілу інтенсивностей для кожного кадру. Наступним етапом проводиться дискретна згортка отриманих масивів з гаусовим ядром. За усередненими розподілами обчислюється взаємна і власна ентропії.

Потім виконується обчислення рахунків, реалізованих у вигляді вкладених циклів, що здійснюють прохід по рядках, стовпцях і диспаратності. Отриманий масив рахунків повертається як результат виконання функції.

В межах тієї ж ітерації виконується виклик іншої зовнішньої функції `stereoSGmap`, яка реалізує обчислення мапи глибини алгоритмом WTA. Робота функції полягає у виконанні циклу, що здійснює прохід по рядкам і стовпцям, в ході якого для кожного пікселя обчислюється мінімальне значення рахунків і повертається його координата на осі диспаратності як значення зсуву для даного пікселя. Результат повертається у вигляді монохромного зображення, в якому яскравість показує значення диспаратності.

Отримана мапа глибини після збільшення розміру вдвічі застосовується як опорна для подальшої ітерації. Обчислення повторюються для зображень розміром $1/8$, $1/4$, $1/2$ від оригінального розміру і на кінцевій ітерації – для вихідних зображень.

Після всіх ітерацій виконується підсумовування рахунків для ліквідації розривів диспаратності і підвищення точності обчислюваної мапи глибини. Підсумовування рахунків представлено зовнішньою функцією `stereoSGcostSum`. На вхід функції надходять розміри кадрів, масив рахунків, коефіцієнти згладження і значення максимальної диспаратності. Повертається результат – масив розмірністю, рівною вхідному масиву рахунків, який містить підсумовані рахунки для кожного пікселя відповідно виразу (3.6).

Здійснення підсумовування починається з ініціалізації масиву рахунків значеннями вхідного масиву для пікселей, що знаходяться на границях зображення. Далі кожен піксель застосовується як відправна точка для методу підсумовування по 8-ми напрямках. Підсумовування здійснюється у зовнішній функції `stereoSGcostTravers`. Вддовж кожного напрямку для кожного пікселя, починаючи з крайнього, виконується підсумовування рахунків за виразом (3.6).

Етап підсумовування рахунків є фінальним для функції `stereoSGfunc`. Наступний етап реалізованого методу – формування мапи глибини за

отриманим масивом підсумовуваних рахунків. Цей крок виконується у функції `stereoSGmap`, описаній вище.

Наступний етап методу – застосування інтерполяції FGS. Реалізація даного методу представлена зовнішньою функцією `fgsWlsFunc`. Алгоритм виконано у вигляді циклу із T ітерацій, в тексті якого відбувається обробка усіх рядків, а потім усіх стовпців зображення. Кількість ітерацій задається користувачем. Обробка рядків здійснюється зовнішньою функцією `fgsWlsH`, а стовпців – `fgsWlsV`. Принцип дії обох функцій схожий, відмінність є тільки в адресації по масиву зображення. Після T ітерацій функція `fgsWlsFunc` дає результат обробки в головну область програми.

Фінальним кроком роботи програми є обчислення точності отриманої мапи глибини (при умові наявності еталонної мапи) та виведення результату на дисплей.

3.4 Обґрунтування вибору тестових зображень

Для вибору тестових зображень в цілях дослідження алгоритму побудови мапи глибини можна зазначити ряд критеріїв:

- для об'єктивного оцінювання точності стереозображення повинно мати еталонну мапу глибини;
- для підтвердження універсальності досліджуваних алгоритмів у тестовому наборі мають бути стереозображення з різним розрізненням;
- для підтвердження надійності у тестовому наборі мають бути стереопари з різним рівнем контрасту фону, в тому числі, з ділянками рівномірного однородного фону.

Очевидно, що для отримання еталону стереовідповідності потрібен підхід, який принципово відрізняється від кореляційно-подібних, які дають велику точність близьку до 100%. Для досягнення цієї точності недостатньо двох зображень, отриманих при природному освітленні. Для збільшення точності 3D реконструкції застосовують алгоритми на основі

активної підсвітки сцени, наприклад за допомогою проектора. Ми не маємо можливості провести подібні експерименти, тому необхідно обрати тестові зображення із загальнодоступних баз, які вже матимуть свій еталон мапи глибини.

Ефективність синтезованих алгоритмів треба досліджувати на різномірних вхідних даних, зокрема на стереозображеннях різного розрізнення і різним рівнем контрасту фону. В роботі для експериментальних перевірок комбінованого алгоритму використано спеціально підібраний набір стереопар.

Автори робіт в галузі встановлення стереовідповідностей частіше за все застосовують «класичний» набір з таких стереопар:

- Tsukuba (384x288),
- Venus (434x383),
- Teddy (450x375),
- Cones (450x375).

Даний набір є головним в найбільш популярній системі рейтингів, широко використовується науковим товариством починаючи з 2002 р. Система рейтингів, набори стереозображень і деякі роботи є частинами відомого порталу комп'ютерного зору. Недоліками обраного набору є мала кількість стереозображень і порівняно невелика роздільна здатність.

Стереозображення з цих наборів показано на рис. 3.2 – рис. 3.4.



Рисунок 3.2 – Стереозображення Teddy



Рисунок 3.3 – Стереозображення Art



Рисунок 3.4 – Стереозображення Моеbius

Крім перелічених стереозображень багато авторів досліджують стереопари: Art (695x555), Books (695x555), Dolls (695x555), Laundry (671x555), Moebius (695x555), Reindeer (671x555). Як зазначають автори досліджень, в даному наборі стереозображення мають великі розміри в пікселях, і на них зображені більш складні сцени для обчислення стереовідповідності.

Важливо зазначити, що усі застосовані стереозображення відкалібровані, що не дозволяє виконати моделювання для етапу ректифікації. Для перевірки цього етапу досліджено пари Umbrella (2960 × 2016) і Classroom1 (3000 × 1920), які мають в наборі некалібровані кадри. Крім того, для даних пар є знімки, отримані за різних умов освітлення, що дозволяє перевірити роботу алгоритму в цих умовах.



Рисунок 3.5 – Стереозображення Umbrella



Рисунок 3.6 – Стереозображення Classroom1

На основі викладеного пропонується використовувати такий набір стереопар для експериментальної перевірки комбінованого алгоритму: Tsukuba, Teddy, Art, Moebius, Umbrella, Classroom1.

3.5 Дослідження комбінованого алгоритму на тестових зображеннях

В даному пункті викладається результат дослідження комбінованого алгоритму на кожному основному кроці алгоритму для стереозображень, включених в тестовий набір. На рис. 3.5 представлено стереозображення Umbrella.

Мапа глибини, отримана до операції підсумовування рахунків показана на рис. 3.7. Помітно, що виникає багато помилок в ототожненні точок і мапа диспаратності виходить сильно розрідженою.



Рисунок 3.7 – Результат етапу SGM до підсумовування рахунків

Алгоритм підсумовування рахунків має виправити дані проблеми і дозволяє отримати зображення, показане на рис. 3.8.



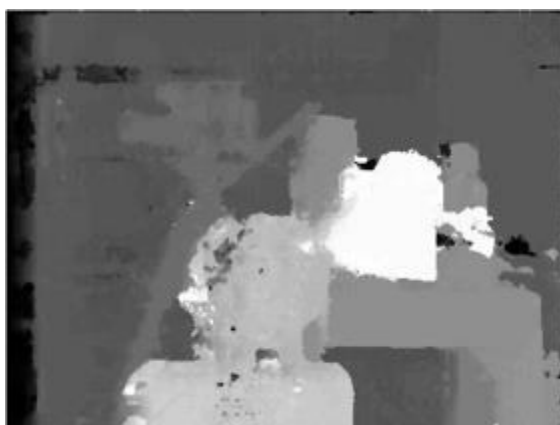
Рисунок 3.8 – Результат етапу SGM і формування мапи глибини алгоритмом WTA

Для усунення залишкового шуму і зіставлення меж мапи глибини з межами вихідного зображення застосовується інтерполяція методом FGS. Результати обробки показані на рис. 3.9.



Рисунок 3.9 – Результат етапу FGS

Отримана мапа глибини є остаточним результатом використання комбінованого алгоритму до стереопари Umbrella. Результати дослідження на інших зображеннях набору показані на рис. 3.10.



а)



б)



в)



г)

Рисунок 3.10 – Мапи глибини по стереозображенням Tsukuba (а), Teddy (б), Art (в), Moebius (г)

3.6 Висновки по розділу 3

В результаті проведення експерименту з застосуванням тестового набору стереозображень було визначено середню точність комбінованого методу. Відсоток помилок склав менше 5%, що означає точність 95%.

Підтверджено, що досліджений комбінований алгоритм успішно справляється із задачею забезпечення підвищеної точності обчислення мапи глибини, дозволяє отримати малу кількість помилок і при цьому має лінійну складність обчислення. Використання методів обробки мапи глибини дещо знижує швидкість обчислень, але це виправдовується суттєвим внеском даних алгоритмів в підвищення точності і гладкості отриманої мапи.

В порівнянні з локальним алгоритмом пошуку алгоритм забезпечує істотно меншу кількість помилок, але це досягається ціною зменшення швидкості роботи. Досліджений комбінований алгоритм лише трохи поступається в точності, але реалізує суттєву перевагу за рахунок меншої складності обчислень.

Якщо порівнювати комбінований алгоритм з класичним напівглобальним алгоритмом, то час роботи збільшився незначно, але підвищення точності отриманої мапи глибини і зіставлення границь мапи глибини з межами вихідного кадру компенсує цей недолік.

ВИСНОВКИ

Технічна довершеність в робототехніці, ігрової індустрії, промисловості і в інших галузях дійшов до такого рівня, що системи комп'ютерного зору, а зокрема стереозору і фотограметрії, набули широкого поширення і швидко розвиваються. Це відбулося у зв'язку з ростом обчислювальних можливостей і високими завданнями, які стоять перед робототехнічними системами і пристроями, адже за допомогою «зору» робот може отримувати великий обсяг інформації про навколишнє середовище.

Окремим випадком застосування систем технічного зору є навігація рухомих робототехнічних систем, будь то автомобілі Tesla, або марсохід Mars Exploration Rover та багато інших.

Системи, технічного зору оснащені 2-ма і більше камерами, тобто системи стереобачення і фотограметрії, дозволяють визначати відстань до об'єкта спостереження. Існує низка проблем стереометоду. Шуми і спотворення на зображення стереопари, відблиски, або протяжні однорідні області в кадрі сильно ускладнюють коректну побудову карти глибини. Тому алгоритм обробки стереозображень повинен мати декілька кроків.

Мета кваліфікаційної роботи – на прикладі реальних зображень стереопар отримати оцінку числових характеристик відомих алгоритмів, що реалізують окремі етапи розрахунку карт глибини, на основі отриманих результатів синтезувати комбінований алгоритм, який зможе забезпечити високу точність карти при незначній обчислювальній складності.

Для досягнення заявленої мети вирішені такі задачі: проведено аналітичний огляд літератури; проведено дослідження відомих алгоритмів побудови мапи глибини; синтезовано алгоритм побудови мапи глибини, який врахує вимоги до швидкодії і точності отриманої мапи глибини; реалізовано синтезований алгоритм у програмному середовищі MATLAB; виконано експериментальне дослідження синтезованого алгоритму; проведено аналіз ефективності синтезованого методу.

В результаті проведеного дослідження встановлено, що найбільший вплив на складність обчислень і точність формування мапи глибини має етап пошуку зв'язаних точок. Саме в цій області ведеться більшість розробок, задачею яких є оптимізація тих чи інших характеристик методів 3D реконструкції.

Дослідження показали, що найкращу точність визначення стереовідповідності забезпечує алгоритм розрізу графа. В той же час він же має найбільшу складність обчислювань.

Метод локального пошуку сильно поступається в точності обчислення 3D моделі, кількість помилок для нього найвища. Незважаючи на те, що своє застосування він отримав завдяки швидкості своєї роботи, ряд методів на даний момент практично не поступається йому, і забезпечують при цьому значно менший відсоток помилок. До таких методів в першу чергу відносяться алгоритм динамічного програмування (оснований на локальному методі) та напівглобальний алгоритм.

Значний вплив на результат можуть мати алгоритми обробки мапи глибини. Їхнє використання не завжди виправдане, тому що ці алгоритми часто збільшують складність обчислювання усього алгоритму. Але при використанні алгоритмів пошуку з порівняно невисокою складністю обчислень можна досягти значущих результатів у збільшенні якості мапи глибини.

Відомі методи ректифікації у своїй більшості мають приблизно однакову швидкість роботи і дають подібні результати. Застосування більш складних методів не виправдано взагалі, оскільки головна мета ректифікації – спростити обчислення на етапі пошуку пов'язаних точок – втрачає сенс при значній обчислювальній складності самого методу ректифікації.

В результаті проведення експерименту з застосуванням тестового набору стереозображень було визначено середню точність комбінованого методу. Відсоток помилок склав менше 5%, що означає точність 95%.

Підтверджено, що досліджений комбінований алгоритм успішно справляється із задачею забезпечення підвищеної точності обчислення мапи глибини, дозволяє отримати малу кількість помилок і при цьому має лінійну складність обчислення. Використання методів обробки мапи глибини дещо знижує швидкість обчислень, але це виправдовується суттєвим внеском даних алгоритмів в підвищення точності і гладкості отриманої мапи.

В порівнянні з локальним алгоритмом пошуку алгоритм забезпечує істотно меншу кількість помилок, але це досягається ціною зменшення швидкості роботи. Досліджений комбінований алгоритм лише трохи поступається в точності, але реалізує суттєву перевагу за рахунок меншої складності обчислень.

Якщо порівнювати комбінований алгоритм з класичним напівглобальним алгоритм, то час роботи збільшився незначно, але підвищення точності отриманої мапи глибини і зіставлення границь мапи глибини з межами вихідного кадру компенсує цей недолік.

ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАНЬ

1. Грузман И.С., Киричук В.С., Косых В.П., Перетягин Г.И., Спектор А.А. Цифровая обработка изображений в информационных системах: Учебное пособие. – Новосибирск: Изд-во НГТУ, 2000. –168 с.
2. Р. Гонсалес, Р. Вудс. Цифровая обработка изображений, Москва: Техносфера, 2005. – 1072 с.
3. Г.Л. Гимельфарб, В.Б. Марченко, В.И. Рыбак. Автоматическое отождествление идентичных точек стереоснимков с учетом неравномерных фотометрических искажений сигналов / Кибернетика. – 1976. – № 4. – С. 107-112.
4. Depth Approximation in Digital Imaging. [Электронный ресурс]. URL: <http://www.r2labs.org/references/depthInImages.pdf>
5. Назаров, А.С. Фотограмметрия: учеб. пособие для студентов вузов. Мн.: ТетраСистемс, 2006. – 368 с.
6. Сойфер А. Методы компьютерной обработки изображений. М., Физматлит, 2003. – 784 с.
7. Schlesinger M.I., Flach B. Some solvable subclass of structural recognition problems. Czech Pattern Recognition Workshop, 2000. – pp. 55-62.
8. Стерео-реконструкция. [Электронный ресурс]. URL: http://courses.graphicon.ru/files/courses/vision/2010/cv_2010_12.pdf
9. Stereo Matching - Local Methods. [Электронный ресурс]. URL: <http://3dstereophoto.blogspot.ru/2012/01/stereo-matching-local-methods.html>
10. Hirschmuller, H. Real-time correlation-based vision with reduced border errors. / Hirschmuller, H., Innocent, P.R., Garibaldi, J.M. // International Journal of Computer Vision. – 2002. – P. 229-246.
11. Disparity map post-filtering. [Электронный ресурс]. URL: http://docs.opencv.org/trunk/d3/d14/tutorial_ximgproc_disparity_filtering.html (дата обращения: 12.04.2017).

12. Uniform Intensity regions. [Электронный ресурс]. URL: <http://ai.stanford.edu/~mitul/cs223b/dp.html> (дата обращения: 12.04.2017).

13. I. Koryttsev, S. Sheiko, V. Kartashov, O. Zubkov, V. Oleynikov, I. Selieznov, M. Anohin. Practical Aspects of Range Determination and Tracking of Small Drones by Their Video Observation // 2020 International Scientific-Practical Conference. Problems of Infocommunications. Science and Technology. Kharkiv, Ukraine. October 6-9, 2020. – 5 p.

14. V. Kartashov, V. Oleynikov, O. Zubkov, S. Sheiko. Optical detection of unmanned air vehicles on a video stream in a real-time // The Fourth International Conference on Information and Telecommunication Technologies and Radio Electronics (UkrMiCo'2019), 9–13 September 2019, Odessa, Ukraine, 4 p.

15. В.М. Карташов, И.В. Корытцев, С.А. Шейко, В.Н. Олейников, О.В. Зубков, С.И. Бабкин. Оптико-электронные методы обнаружения воздушных объектов и измерения их координат // Радиотехника: Всеукр. межвед. науч.-техн. сб. – 2020. – Вып. 202. – С. 153 – 159.

16. О.В. Зубков, С.А. Шейко, В.Н. Олейников, В.М. Карташов, И.В. Корытцев, С.И. Бабкин. Исследование эффективности детектирования и распознавания изображений дронов по видеопотоку // Радиотехника: Всеукр. межвед. науч.-техн. сб. – 2020. – Вып. 202. – С. 136 – 146.

17. И.В. Корытцев, С.А. Шейко, В.М. Карташов, О.В. Зубков, В.Н. Олейников, С.И. Бабкин, И.С. Селезнев. Обработка сигналов при пеленгации и определении дальности до малоразмерных БПЛА в оптическом и инфракрасном диапазонах // Радиотехника: Всеукр. межвед. науч.-техн. сб. – 2020. – Вып. 202. – С. 125 – 134.

18. В.М. Карташов, В.Н. Олейников, М.М. Колендовская, Л.П. Тимошенко, Н.В. Рыбников, А.И. Капуста. Комплексирование изображений при обнаружении беспилотных летательных аппаратов // Радиотехника: Всеукр. межвед. науч.-техн. сб. – 2020. – Вып. 201. – С. 120 – 129.

19. Machine Vision and Navigation / Editors: Sergiyenko, Oleg, Flores-Fuentes, Wendy, Mercorelli, Paolo. Springer, Cham.

20. Ivanov M., Sergiyenko O., Tyrsa V., Mercorelli P., Kartashov V.M., Hernandez W., Sheiko S., Kolendovska M. Individual Scans Fusion in Virtual Knowledge Base for Navigation of Mobile Robotic Group with 3D TVS // IECON 2018 - 44th Annual Conference of the IEEE Industrial Electronics Society. – P. 3187–3192.

21. В.Н. Олейников, О.В. Зубков, В.М. Карташов, И.В. Корытцев, С.И. Бабкин, С.А. Шейко, И.С. Селезнев. Экспериментальная оценка эффективности алгоритмов пеленгования беспилотных летательных аппаратов по акустическому излучению. Радиотехника: Всеукр. міжвед. науч.-техн. сб. – 2019. – Вып. 199. – С. 29 – 37.

22. V. Kartashov, V. Oleynikov, I. Koryttsev, S. Sheiko, O. Zubkov, S. Babkin, I. Selieznov. Use of Acoustic Signature for Detection, Recognition and Direction Finding of Small Unmanned Aerial Vehicles. 2020 IEEE 15th International Conference on Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering (TCSET). 2020. 4 p.

23. Kartashov V.M., Oleynikov V.N, Zubkov O.V., Koryttsev I.V., Babkin S. I., Sheiko S.A., Kolendovskaya M.M. Spatial-temporal Processing of acoustic Signals of Unmanned Aerial Vehicles/ Telecommunications and Radio Engineering. –New York. – 2020. – Vol. 79, №9. – P.769-780.

24. V. Kartashov, V. Oleynikov , I. Koryttsev, S. Sheiko, O. Zubkov, S. Babkin. Processing of Wide Band Acoustic Signals During Detection of Unmanned Aerial Vehicles // 2020 IEEE Ukrainian Microwave Week (UkrMW). Kharkiv, Ukraine, September 21 - 25, 2020. Volume 1 on 2020 IEEE 12th International Conference on Antenna Theory and Techniques (ICATT). pp. 35-39.

25. Hartley R.I. Theory and Practice of Projective Rectification. International Journal of Computer Vision. – Vol. 35. – 1999. – P. 115–127.

26. Mallon, J. Projective Rectification from the Fundamental Matrix. / J. Mallon, P. Whelan // Image and Vision Computing. – 2005. – Vol. 23. – Issue 7. – P. 643-650.

27. Ndhlovu T. An investigation into stereo algorithms: An emphasis on local-matching. / University of Cape Town, 2011, 75 p.

28. Birchfield, S., Tomasi, C. A pixel dissimilarity measure that is insensitive to image sampling. / IEEE Transactions on Pattern Analysis and Machine Intelligence. - 1998. - Vol. 20. - № 4. - P. 401–406.
29. Hirschmuller, H., Scharstein, D. Evaluation of stereo matching costs on images with radiometric differences. / IEEE Transactions on Pattern Analysis and Machine Intelligence. - 2009. - Vol. 31. - № 9. - P. 1582–1599. 26
30. A. L.Yuille, T. Poggio, “A generalized ordering constraint for stereo correspondence,” MIT, A.I.Memo 777, 1984, aI Lab. 7, 27.
31. Вахитов А.Т., Гуревич Л.С., Павленко Д.В. Обзор алгоритмов стереозрения. [Электронный ресурс]. URL: <http://www.math.spbu.ru/user/gran/soi4/pavlenko.pdf>.
32. Tomasi C. Global Stereo in Polynomial Time [Электронный ресурс]. URL: <https://users.cs.duke.edu/~tomasi/papers/tomasi/tomasiYork07.pdf>
33. H. Hirschmuller. Accurate and efficient stereo processing by semi-global matching and mutual information. IEEE Conference on Computer Vision and Pattern Recognition, 2006.
34. S. Birchfield, C. Tomasi. Depth discontinuities by pixel-to-pixel stereo. In Proceedings of the Sixth IEEE International Conference on Computer Vision, pages 1073–1080, Mumbai, India, January 1998.
35. E. Riegler, G. E. Kinkelund, C. N. Manchón, . Badiu, et al. Merging Belief Propagation and the Mean Field Approximation: A Free Energy Approach [Электронный ресурс]. URL: <https://arxiv.org/pdf/1112.0467.pdf> .
36. Y. Boykov, O. Veksler, and R. Zabih. Efficient approximate energy minimization via graph cuts. IEEE Transactions on Pattern Analysis and Machine Intelligence, 23(11):1222–1239, 2001.
37. G. Egnal and R. P. Wildes, “Detecting binocular half-occlusions: Empirical comparisons of five approaches,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 24, no. 8, pp. 1127–1133, 2002.
38. K. Konolige, “Small vision systems: hardware and implementation,” in Eighth International Symposium on Robotics Research, 1997, p. 111–116.

39. J. J. Little and W. E. Gillett, "Direct evidence for occlusion in stereo and motion," in Proceedings of the First European Conference on Computer Vision. London, UK: Springer-Verlag, 1990, pp. 336–340.
40. A. Luo and H. Burkhardt, "An intensity-based cooperative bidirectional stereo matching with simultaneous detection of discontinuities and occlusions," International Journal of Computer Vision, vol. 15, no. 3, pp. 171–188, 1995.
41. Q. Yang, R. Yang, J. Davis, and D. Nister, "Spatial-depth super resolution for range images," IEEE Conference on Computer Vision and Pattern Recognition, pp. 1–8, 2007.
42. Wöhler C. 3D Computer Vision. Efficient Methods and Applications. Berlin: Springer-Verlag Berlin Heidelberg, 2009. – 389 p.
43. Бабенко В.С. Оптика телевизионных устройств. М.: Радио и связь, 1982. – 256 с.
44. Методичні вказівки з виконання атестаційної магістерської роботи за спеціальністю 8.05090102 «Апаратура радіозв'язку, радіомовлення і телебачення». Освітньо-кваліфікаційний рівень – магістр / Упоряд. В.М. Карташов, В.А. Тихонов, І.В. Савченко – Харків: ХНУРЕ, 2012. – 68 с.