

НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ «ОДЕСЬКА ПОЛІТЕХНІКА»  
МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ

НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ «ОДЕСЬКА ПОЛІТЕХНІКА»  
МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ

*Кваліфікаційна наукова  
праця на правах рукопису*

**КОНДРАТЬЄВ СЕРГІЙ БОРИСОВИЧ**

УДК 004.93

**ДИСЕРТАЦІЯ**

**МЕТОДИ УДОСКОНАЛЕННЯ ПОБУДОВИ КАРТИ ГЛИБИН ДЛЯ СИСТЕМ  
ТЕХНІЧНОГО ЗОРУ**

Спеціальність: 122 – «Комп'ютерні науки»

Галузь знань: 12 – «Інформаційні технології»

Подається на здобуття наукового ступеня доктора філософії

Дисертація містить результати власних досліджень. Використання ідей, результатів і текстів інших авторів мають посилання на відповідне джерело.

  
\_\_\_\_\_ С. Б. Кондратьєв

Науковий керівник:

Антощук Світлана Григорівна, доктор технічних наук, професор

Одеса – 2025

## АНОТАЦІЯ

*Кондратьєв С.Б.* Методи удосконалення карти глибин для систем технічного зору. – Кваліфікаційна наукова праця на правах рукопису.

Дисертація на здобуття наукового ступеня доктора філософії за спеціальністю 122 – Комп'ютерні науки. – Національний університет «Одеська політехніка», МОН України, Одеса, 2025.

У **вступі** обґрунтовано актуальність роботи, що пов'язана з розробкою методів удосконалення побудови карти глибин для систем технічного зору. Визначено об'єкт, предмет, мету, задачі та методи дослідження; наведено наукову новизну та практичне значення отриманих результатів; висвітлено особистий внесок здобувача, наведено основні дані щодо апробації роботи, публікацій, відомості про структуру та загальну характеристику дисертаційної роботи.

В **першому розділі** дисертаційної роботи проведено аналіз проблеми побудови карти глибин для систем технічного зору.

Проаналізовано основне призначення та принципи побудови систем технічного зору, розглянуті види систем та задачі, які вирішують такі системи. Визначено, що задача оцінки глибини сцени є ключовим компонентом систем технічного зору, що забезпечує тривимірне сприйняття навколишнього середовища. Це дозволяє стверджувати, що однією з перспективних тем досліджень в області побудови систем технічного зору є удосконалення методів побудови карти глибин.

Визначено, що методи побудови карти глибин можна розділити на активні та пасивні. Встановлені переваги та недоліки кожного типу методів, а також визначені прикладні ситуації, в яких доцільніше застосовувати ті чи інші методи побудови карт глибини.

Детальний аналіз пасивних методів дозволив визначити, що ці методи також поділяються на локальні та глобальні методи в залежності від принципів роботи алгоритму стерео зіставлення, на якому вони базуються. Найбільш перспективним підходом пасивного отримання карти глибин вважаються локальні методи на

основі адаптивної ваги. Незважаючи на відносно задовільні результати, які дають методи на основі адаптивної ваги, попіксельне обчислення ваги та оцінка глибини є дуже трудомісткими завданнями, тобто вони мають високу обчислювальну надмірність. Для усунення цього недоліка та підвищення швидкості побудови карти глибини було запропоновано пасивний локальний метод побудови карти глибини на основі адаптивної ваги з використанням процедури адаптивного зіставлення та афінно-інваріантних ознак.

Також подальший аналіз пасивних методів отримання карти глибини показав, що карти глибини для динамічної оцінки навколишнього середовища будуються покадрово, що призводить до появи артефактів, внаслідок того, що значення глибини статичних об'єктів повинні залишатися однаковими у сусідніх кадрах, але часто вони оцінюються як різні значення (ця проблема часто називається часовою неузгодженістю). Такі помилки значно погіршують візуальну якість синтезованого представлення, а також ефективність кодування послідовностей карт глибин. Для подолання цього недоліку було запропоновано метод покращення часової узгодженості для послідовностей карт глибини на базі показника надійності глибини та динамічних атрибутів регіонів.

Аналіз активних методів отримання карти глибин показав, що такі методи часто продукують карти глибин з низькою роздільною здатністю, зашумленими значеннями глибини та прогалинами на зображенні карти, тому отримані карти глибини покращуються за допомогою наявного вихідного кольорового зображення (такі методи покращення називаються керованими). Однак, існуючі керовані методи покращення карти глибини не оцінюють в явному вигляді неузгодженість країв між кольоровим зображенням і отриманою картою глибини, тому вони не можуть адаптивно керувати процесом зіставлення кольорового зображення під час покращення карти глибин. Це призводить до появи артефактів копіювання текстури та розмиття країв на карті глибини. Для подолання цього недоліку було запропоновано керовані методи покращення карти глибини на базі марковського випадкового поля, з використанням процедури оцінки неузгодженості країв

вихідного кольорового зображення та отриманої карти глибин з жорстким та м'яким прийняттям рішень.

Крім того аналіз керованих методів покращення карти глибини на базі кольорового зображення показав, що спорідненість країв кольорового зображення та карти глибин зазвичай обчислюється лише на основі різниці в кольорі та глибині між пікселем та його сусідніми пікселями на кольоровому зображенні та карті глибини, відповідно. Така схема обчислень ігнорує локальну структуру карти глибини. Тому, у випадку великого коефіцієнта дискретизації карти глибини, підсумкова карта буде мати надмірно згладжені краї глибин. Для подолання цього недоліку та збереження структури на картах глибин дуже низької якості було запропоновано керований метод покращення карти глибини низької якості на базі марковського випадкового поля, з використанням процедури оцінки неузгодженості країв з м'яким прийняттям рішень та представленням локальної структури карти глибин з використанням мінімального кістякового лісу.

За результатами проведеного аналізу сформульовано мету та завдання досліджень.

В **другому розділі** дисертаційної роботи запропоновані підходи до удосконалення пасивних методів побудови карт глибини.

**Для підвищення швидкості пасивних методів побудови карти глибини за допомогою множини кольорових зображень** було запропоновано пасивний локальний метод побудови карти глибини на основі адаптивної ваги з використанням процедури адаптивного зіставлення та афінно-інваріантних ознак.

Запропонований метод складається з двох етапів:

1. Початкова оцінка глибини. Для її отримання виконуються наступні кроки: агрегація витрат, яка виконується за допомогою запропонованої процедури адаптивного зіставлення; уточнення витрат для гладких регіонів; побудова карти глибини за допомогою підходу «переможець отримує все».

2. Уточнення карти глибини (корекція карти для гладких та негладких регіонів). Для корекції карти для гладких регіонів виконується пошук в околі кожної ненадійної гладкої області, а всі надійні гладкі області заносяться до списку

кандидатів. Якщо середнє значення колірних каналів RGB області у списку значно відрізняється від середнього значення поточної ненадійної згладженої області, то область буде видалено зі списку. Далі відбувається оновлення глибини області у списку, колір якої є найближчим до неї. Для негладких регіонів, значення їхньої глибини уточняється за допомогою багатосторонньої фільтрації. У процесі фільтрації ваги пікселів, які належать до ненадійних згладжених областей та ненадійних пікселів у згладжених областях, прирівнюються до нуля.

Експериментальні дослідження на різних наборах даних показали зменшення часу на побудову карти глибини на 19,9% – 24,1% у порівнянні з методом-аналогом.

**Сформульовано перший пункт наукової новизни:** *удосконалено пасивний локальний метод побудови карти глибини за рахунок використання процедури адаптивного зіставлення та афінно-інваріантних ознак, що дозволило підвищити швидкість побудови карти глибини.*

**Для зменшення кількості артефактів при побудові віртуального зображення з використанням послідовних карт глибини,** було запропоновано метод покращення часової узгодженості для послідовностей карт глибини за рахунок впровадження показника надійності глибини та використання динамічних атрибутів регіонів.

Запропонований метод складається з двох етапів:

1. Оцінка вірогідності того, що певна область зображення належить до статичного регіону. Базуючись на спостереженні, що статичні області мають однаковий колір між сусідніми кадрами, для оцінки того, чи належать пікселі до статичних областей, використовується значення середнього абсолютного відхилення. Оцінювання виконується у межах локального вікна з центром у пікселі, який оцінюється.

2. Оцінка надійності значень глибини – через наявність шуму в оціночних картах глибин, достовірність значень глибин слід враховувати в явному вигляді, що дає більшу вагу більш достовірній глибині. Для кожного пікселя на карті глибини поточного ракурсу, надійність глибини можна оцінити через значення середнього

абсолютного відхилення між локальними вікнами з центром у такій парі пікселів. Для врахування оклюзії оцінка виконується на сусідніх видах зліва та справа, при цьому вибирається менше значення середнього абсолютного відхилення.

Проведені експериментальні дослідження на різних наборах даних, які показали збільшення часової узгодженості на 10,3% – 44,75% та покращення ефективності кодування карти глибини на 13,3% – 51,6% у порівнянні з розглянутим методом-аналогом.

**Сформульовано другий пункт наукової новизни:** *удосконалено* метод покращення часової узгодженості для послідовностей карт глибини за рахунок урахування показника надійності глибини та використання динамічних атрибутів регіонів, що дозволило покращити часову узгодженість та покращити ефективність кодування карти глибини.

**Для підвищення точності та зменшення обчислювальних витрат при побудові карти глибин** запропоновано модифікований метод StereoBM-Mod з вейвлет-попередньою обробкою. Метод дозволяє локалізувати інформативні області та звужити зону пошуку відповідностей. Експериментально доведено, що StereoBM-Mod забезпечує вищу точність (на 1.5–2.4 дБ за PSNR) порівняно з класичним StereoBM, зберігаючи низьке енергоспоживання та придатність до роботи в реальному часі. Він перевершує складніші методи за балансом точності, швидкодії та енергоефективності. Метод рекомендовано для використання у вбудованих та мобільних системах технічного зору.

**Сформульовано третій пункт наукової новизни:** *набув подальший розвиток* метод локального стереозіставлення StereoBM шляхом інтеграції наближеного вейвлет-перетворення Гаара для попереднього виділення контурних ознак та локалізації екстремумів. Використання запропонованого методу дозволило зменшити обчислювальну складність процесу побудови карти глибини, зберігаючи при цьому точність у слабо текстурованих зонах.

В **третьому розділі** дисертаційної роботи запропоновані підходи до удосконалення карт глибини, що отримані за допомогою активних методів.

**Для покращення точності карти глибини, що побудована за допомогою керованих методів покращення карти глибини з використанням активних методів, запропоновано керований метод покращення точності карти глибини на базі марковського випадкового поля за рахунок впровадження процедури оцінки неузгодженості країв вихідного кольорового зображення та отриманої карти глибини з жорстким та м'яким прийняттям рішень.**

У якості базового методу використовується метод покращення точності карти глибини на базі марковського випадкового поля, до якого додається процедура оцінки неузгодженості країв між кольоровим зображенням та отриманої активним методом карти глибини.

Неузгоджені краї можна розділити на два типи: перші одержані внаслідок грубої інтерполяції, але можуть бути уточнені за допомогою наведення на високоякісні кольорові границі, а другі – внаслідок неузгодженості країв. При використанні жорсткого прийняття рішення, відбувається класифікація цих двох випадків. Виходячи з того, що кількість помилок має бути меншою за можливості кодування з виправленням помилок, припускається, що глибинний край погіршується грубою інтерполяцією, коли зміщення між ним і найближчим кольоровим краєм є меншим за порогове значення. Інший випадок трактується як справжня неузгодженість країв.

Хоча підхід на базі жорсткого прийняття рішення дає задовільні результати для малих коефіцієнтів дискретизації, продуктивність методу суттєво погіршується для великих коефіцієнтів дискретизації. Причина полягає в тому, що вимірювання неузгодженості країв з жорсткою дискретизацією не може добре працювати, коли якість грубо інтерпольованої карти глибини дуже низька. Таким чином, пропонується підхід на базі м'якого прийняття рішень, який більш точно вимірює ступінь неузгодженості, що виникає між глибинними краями та відповідними кольоровими краями у числовому вигляді.

Експериментальні дослідження запропонованого методу на різних наборах даних показали збільшення точності підсумкової карти глибини на 4,3% – 22,19% для коефіцієнтів інтерполяції 2 – 8 для підходу з жорстким прийняттям рішень, та

збільшення точності підсумкової карти глибини на 5,4% – 7,7% для коефіцієнтів інтерполяції 8 – 16 для підходу з м'яким прийняттям рішень.

**Сформульовано четвертий пункт наукової новизни:** *удосконалено* керований метод покращення точності карти глибини на базі марковського випадкового поля за рахунок впровадження процедури оцінки неузгодженості країв вихідного кольорового зображення та отриманої карти глибини з жорстким та м'яким прийняттям рішень, що дозволило покращити точність карти глибини.

**Для покращення точності карти глибини, побудованої за допомогою керованих методів покращення карти глибини, яка отримана активними методами, з великим коефіцієнтом дискретизації** було запропоновано керований метод покращення точності карти глибини низької якості на базі марковського випадкового поля за рахунок впровадження процедури оцінки неузгодженості країв з м'яким прийняттям рішень та представленням локальної структури карти глибин з використанням мінімального кістякового лісу.

Запропонований метод використовує мінімальний кістяковий ліс у якості фільтру, який використовується для автоматичного видалення двох несхожих пікселів, які знаходяться близько один до одного в просторовій області, що робить відстань між деревами метрикою, яка враховує ребра. Для запобігання проблеми «протікання країв» запропонований метод будує мінімальний кістяковий ліс для кожного суперпікселя на кольоровому зображенні, згенерованому надмірною сегментацією.

Експериментальні дослідження запропонованого методу на різних наборах даних показали збільшення точності підсумкової карти глибини на 2,8% – 11,87% для коефіцієнтів інтерполяції 8 – 16 у порівнянні з найбільш ефективними методами-аналогами, включаючи запропонований раніше метод з м'яким прийняттям рішень.

**Сформульовано п'ятий пункт наукової новизни:** *удосконалено* керований метод покращення точності карти глибини низької якості на базі марковського випадкового поля за рахунок впровадження процедури оцінки неузгодженості країв з м'яким прийняттям рішень та представленням локальної структури карти

глибин з використанням мінімального кістякового лісу, що дозволило покращити точність карти глибини з великим коефіцієнтом дискретизації.

В четвертому розділі дисертаційної роботи проведено апробацію запропонованих методів побудови карти глибини у складі стереозорової підсистеми мобільної робототехнічної платформи. На основі розробленого експериментального стенду, що включає одноплатний комп'ютер, модуль камери та ToF-сенсор, вдалося реалізувати повний цикл глибинної реконструкції – від обробки вхідних зображень до отримання карти глибини в реальному часі. Це забезпечило можливість перевірити ефективність розроблених методів не лише в теоретичному чи лабораторному контексті, а й в умовах, наближених до практичного застосування в задачах автономної навігації.

Апробація проводилася на даних, отриманих як зі стандартних наборів, так і з реального сенсорного обладнання, що дало змогу оцінити роботу алгоритмів за критеріями точності, стабільності, енергоспоживання та обчислювальної складності. Усі методи були реалізовані з урахуванням обмежень вбудованих систем, зокрема щодо використання апаратних ресурсів, енергоефективності та вимог до обробки в реальному часі.

Загалом, результати апробації свідчать про практичну придатність запропонованих методів до впровадження в системи реального часу. Висновки, отримані на основі експериментальних вимірювань та порівнянь із базовими підходами, підтверджують доцільність їх використання в системах технічного зору.

*Ключові слова:* карти глибин, системи технічного зору, комп'ютерний зір, обробка зображень, просторовий аналіз, 3D-реконструкція, стереобачення, сенсори глибини, оцінка глибини, програмне забезпечення, представлення та аналіз сцен, покращення ефективності, позиціонування об'єктів, виділення контурів.

## **СПИСОК ОПУБЛІКОВАНИХ ПРАЦЬ ЗА ТЕМОЮ ДИСЕРТАЦІЇ**

*Наукові праці, в яких опубліковані основні наукові результати дисертації:*

1. Arsirii, O.O., Yadrova, M.V., Kondratyev, S.B., & Stelmakh, D.E. Development of the intelligent software and hardware subsystem for capturing an object by robot manipulator. Herald of Advanced Information Technology. 2020; Том 3(2) 42–51. doi.org/10.15276/hait.02.2020.4. Видання включено до переліку наукових фахових видань України, категорія «Б».

<https://hait.od.ua/index.php/journal/article/view/35>

2. Кондратьєв С.Б., Костенко В.Л., Ядро́ва М.В. Метод контурів для позиціонування об'єктів в мобільних системах комп'ютерного зору. Вісник Національного технічного університету «ХПІ». 2021. № 2(8). С. 62–69. doi.org/10.20998/2413-4295.2021.02.09. Видання включено до переліку наукових фахових видань України, категорія «Б».

<https://repository.kpi.kharkov.ua/items/2e80004e-f419-4194-b71b-5d68aea4424a>

3. Antoshchuk, S.G., Kondratyev, S.B., Shcherbakova, G.Y., & Hodovychenko, M.A. Depth map generation for mobile navigation systems based on objects localization in images. Herald of Advanced Information Technology. 2022. Том 5(1). С. 11–18. doi.org/10.15276/hait.05.2022.1. Видання включено до переліку наукових фахових видань України, категорія «Б».

<https://hait.od.ua/index.php/journal/article/view/128>

4. Kondratyev, S.B., Antoshchuk, S.G. & Hodovychenko, M.A. Methods for refining the depth map obtained from depth sensors. Applied Aspects of Information Technology. 2024. № 7(4). С. 336–347. doi.org/10.15276/hait.05.2022.1. Видання включено до переліку наукових фахових видань України, категорія «Б».

<https://aait.od.ua/index.php/journal/article/view/261>

5. Кондратьєв С.Б. Локальний метод побудови карти глибин на базі адаптивної схеми зіставлення та афінних інваріантних ознак. Наука і техніка сьогодні. 2024; № 12(40). С. 1245–1258. doi.org/10.52058/2786-6025-2024-12(40)-1245-1258. Видання включено до переліку наукових фахових видань України, категорія «Б».

<http://perspectives.pp.ua/index.php/nts/article/view/17042>

6. Kondratyev, S.B., Antoshchuk, S.G., & Hodovychenko, M.A. Robust depth map refining using color image. Herald of Advanced Information Technology. 2024; № 7(4): С. 361–370. doi.org/10.15276/hait.07.2024.25. *Видання включено до переліку наукових фахових видань України, категорія «Б».*

<https://hait.od.ua/index.php/journal/article/view/232>

*Наукові праці, які засвідчують апробацію матеріалів дисертації:*

7. Shcherbakova, G., Antoshchuk, S., Koshutina, D., Sakhno, K., Kondratiev, S. (2025). Adaptive Clustering for Distribution Parameter Estimation in Technical Diagnostics. In: Dovgyi, S., Siemens, E., Globa, L., Kopyika, O., Stryzhak, O. (eds) Applied Innovations in Information and Communication Technology. ICAIT 2024. Lecture Notes in Networks and Systems, vol 1338. Springer, Cham. [doi.org/10.1007/978-3-031-89296-7\\_19](https://doi.org/10.1007/978-3-031-89296-7_19). *Видання проіндексовано у базі даних Scopus.*

[https://icait.org/proceedings/12th\\_ICAIT\\_1/2\\_8\\_ICAIT\\_2024\\_Part\\_2\\_paper\\_6.pdf](https://icait.org/proceedings/12th_ICAIT_1/2_8_ICAIT_2024_Part_2_paper_6.pdf)

8. Antoshchuk S., Shcherbakova G., Kondratyev S., Koshutina, D., Usov O. Wavelet transform based optimization method for Three-Dimensional computer vision // Proceedings of the 12th International Conference Information Control Systems & Technologies (ICST 2024), Odesa, Ukraine, September 23–25, 2024, pp. 471–482. *Видання проіндексовано у базі даних Scopus.*

<https://ceur-ws.org/Vol-3790/paper41.pdf>

9. Kondratiev S., Koshutina D., Liubomska O., Baskov I. Research on Energy Efficiency of Wi-Fi IoT Systems on Renesas DA16200 Platform // Proceedings of International Conference on Applied Innovation in IT, Koethen, Germany, March 7, 2024, 12(1), pp. 143–148. *Видання проіндексовано у базі даних Scopus.*

[https://icait.org/proceedings/12th\\_ICAIT\\_1/2\\_11\\_ICAIT\\_2024\\_Part\\_2\\_paper\\_9.pdf](https://icait.org/proceedings/12th_ICAIT_1/2_11_ICAIT_2024_Part_2_paper_9.pdf)

10. Антошук С.Г., Щербакова Г.Ю., Кондратьев С.Б., Кошутіна Д.В. Наближений метод оцінки глибини об'єктів на основі вейвлет-перетворення. Матеріали X Міжнародної науково-практичної конференції Інформатика.

Культура. Техніка. 2024 25–27 вересня 2024 р. м. Одеса (Україна). 2024; 1(1): 197–201.

<https://ics60.aait.od.ua/zbirnik2024.pdf>

11. Світлана Антощук, Галина Щербакова, *Сергій Кондратьєв*, Дар'я Кошутіна, Олександр Усов. Метод оптимізації на основі вейвлет-перетворення для наближених методів оцінки глибини // Інформаційні управляючі системи та технології (ІУСТ-Одеса-2024): Матеріали XII Міжнародної науково-практичної конференції, 23–25 вересня 2024 р. м. Одеса (Україна). вип. ред. Вичужанін. – Одеса: Видавничий дім «Гельветика», 2024. – 334 с.123–127

<http://icst-conf.com/2024.pdf>

12. Антощук С.Г., Арсірій О.О., *Кондратьєв С.Б.* та ін. Гібридні моделі представлення та методи обробки зображень в просторі вейвлет-перетворення. 2020. 160 с.

<https://discovery.kpi.ua/Record/000634213>

13. Антощук С.Г., Арсірій О.О., *Кондратьєв С.Б.* та ін. Аналіз та розпізнавання зображень у просторі вейвлет-перетворення. 2022. 126 с.

<http://www.irbis-nbuv.gov.ua/publ/REF-0000823682>

## ABSTRACT

*Kondratiev S.B.* Methods for improvement of depth map for technical vision systems. – Qualification scientific work in the form of manuscript.

Thesis for the PhD degree in specialty 122 Computer science. – Odessa Polytechnic National University, Ministry of Education and Science of Ukraine, Odesa, 2025.

The **introduction** shows the relevance of using a depth map for technical vision systems. Particular attention is paid to the process of building a depth map. The object, subject, tasks and methods of the study are defined; the scientific novelty and practical significance of the results are presented; the personal contribution of the applicant is highlighted.

The **first chapter** of the thesis analyzes the problem of building a depth map for technical vision systems.

The main purpose and principles of building technical vision systems are analyzed, the types of systems and tasks that such systems solve are considered. It is determined that the task of estimating the depth of a scene is a key component of technical vision systems that provides three-dimensional perception of the environment. This allows us to assert that one of the promising research topics in the field of building technical vision systems is the improvement of methods for building depth maps.

The task of constructing a depth map is analyzed. It is determined that depth mapping methods can be divided into active and passive methods. The advantages and disadvantages of each type of method are determined, and the situations in which it is more expedient to use certain methods of depth mapping are identified.

A detailed analysis of passive methods has made it possible to determine that these methods are also divided into local and global methods, depending on the principles of the stereo matching algorithm on which these methods are based. Local methods based on adaptive weight are considered to be the most promising approach to passively obtaining a depth map. Despite the relatively satisfactory results provided by adaptive weight-based methods, pixel-by-pixel weight calculation and depth estimation are very

time-consuming tasks. They have a high computational redundancy because depth maps are always smooth.

To overcome these shortcomings and increase the speed of depth map construction, a passive local method of depth map construction based on responsive weight using an ad hoc matching procedure and affiliative invertible features was proposed.

Further analysis of passive depth mapping methods has shown that depth maps are constructed frame by frame for dynamic environmental assessment. Frame-by-frame construction of depth maps often leads to artifacts, due to the fact that the depth values of static objects should remain the same in subsequent frames, but often they are valued as separate values (this problem is often called transient mismatch). Such errors significantly degrade the visual quality of the synthesized representation, as well as the efficiency of encoding depth map sequences.

To overcome this drawback, a method for improving temporary coherence for depth map sequences based on the depth accuracy index and dynamic features of districts was proposed.

An analysis of active depth mapping methods has shown that such methods often produce depth maps with low resolution, noisy depth values, and gaps in the map image, so the resulting depth maps are enhanced using the available source color image (such enhancement methods are called supervised methods). However, existing supervised depth map enhancement methods do not explicitly evaluate the edge mismatch between the color image and the resulting depth map, so they cannot adaptively control the color image matching process during depth map enhancement. This results in texture copying artifacts and edge blurring in the depth map.

To overcome this disadvantage, we propose controlled methods for improving the depth map based on a Markov random field, using a procedure for assessing the mismatch between the edges of the original color image and the resulting depth map with hard and soft problem solving.

A further analysis of supervised methods for color image-based depth map enhancement revealed that the edge similarity of a color image and depth map is usually calculated only based on the difference in color and depth between a pixel and its

neighboring pixels in the color image and depth map, respectively. This calculation scheme ignores the local organization of the depth map. Therefore, in the case of a large depth map sampling ratio, the final map will have excessively smoothed depth edges.

To overcome this drawback and conserve the structure in very low quality depth maps, a supervised method for improving a low quality depth map based on a Markov random field, using an edge mismatch scoring procedure with soft solution and representation of the local organization of the depth map using a sparse skeleton forest, was proposed.

Based on the results of the analysis, the goal and objectives of the research are formulated.

The **second chapter** of the thesis proposes approaches to improving passive depth mapping methods.

**To improve the speed of passive depth mapping methods using a set of color images**, a passive local depth mapping method based on responsive weight using an ad hoc matching procedure and affiliative invertible attributes was proposed.

The proposed method consists of two stages:

1. Primary depth assessment – at this stage, the primary depth map is obtained. The first step of this stage is cost estimation, which is performed using the proposed predictive mapping scheme, followed by cost adjustment for smooth regions, and then depth map construction using the winner-take-all approach.

2. Correcting the depth map – at this stage, the map is corrected for smooth and non-smooth regions. To correct the map for smooth regions, a search is performed in the vicinity of each not reliable smooth region, and all trustworthy smooth regions are added to the candidate list. If the average value of the RGB color channels of a region in the list differs significantly from the average value of the current non-trusted smooth region, the region is removed from the list. The next step is to update the depth of the region in the list with the color closest to it. For non-smooth regions, their depth is updated using multipath filtering. In the process of filtering, the weights of pixels assigned to non-reliable smoothed regions and non-reliable pixels in smoothed regions are set to zero.

Experimental studies on different datasets have shown a 19,9% to 24,1% reduction in the time required to build a depth map compared to the analog method.

**The first point of scientific novelty is formulated:** the passive local method of depth mapping *was improved* by using an adequate matching procedure and affiliative indifferent attributes. The use of the proposed method made it possible to increase the speed of depth map construction.

**To decrease the number of artifacts in the construction of a virtual image using sequential depth maps,** a method for improving transient stability for depth map sequences was proposed by introducing a stability measure and using a dynamic feature of areas. Using the proposed method allows to improve timing stability and improve the efficiency of depth map compression.

The proposed method consists of two stages:

1. Estimation of the probability that a certain segment of the image belongs to a stationary domain - based on the observation that stationary domains have the same color between neighboring frames, the average absolute difference is used to estimate whether pixels belong to stationary domains. The estimation is performed within a local window centered on the pixel being estimated;

2. Assessing the confidence of depth scores – due to the presence of noise in the estimated depth maps, the confidence of depth scores should be explicitly considered, giving more weight to more reliable depths. For each pixel in the depth map of the current view, the reliability of the depth can be estimated by the average of the absolute difference between the local windows centered on such a pair of pixels. To take into account occlusion, the assessment is performed on neighboring views to the left and right, and the smaller average value of the absolute difference is selected.

Experimental studies on different datasets have shown an increase in temporal consistency by 10.3% - 44.75% compared to the analog method and an improvement in depth map encoding efficiency by 13.3% - 51.6% compared to the analog method under consideration.

**The second point of scientific novelty is formulated:** a method for improving timing coherence for depth map sequences *was improved* by introducing a measure of the

depth confidence and using regional dynamic features. The use of the proposed method allowed to improve the time coherence and improve the efficiency of depth map compression.

**To improve the accuracy and reduce computational costs of building a depth map**, a modified StereoBM-Mod method with wavelet preprocessing is proposed. The method allows localizing informative areas and narrowing the search area for correspondences. It has been experimentally proven that StereoBM-Mod provides higher accuracy (by 1.5-2.4 dB in terms of PSNR) compared to the classical StereoBM, while maintaining low power consumption and suitability for real-time operation. It outperforms more complex methods in terms of the balance of accuracy, performance, and energy efficiency. The method is recommended for use in embedded and mobile vision systems.

**The third point of scientific novelty was formulated:** the method of local stereo matching StereoBM was further developed by integrating the approximate wavelet transform of Gaar for preliminary selection of contour features and localization of extremes. The use of the proposed method made it possible to reduce the computational complexity of the depth mapping process while maintaining accuracy in weakly textured areas.

In the **third chapter** of the thesis, approaches to improving depth maps obtained using active methods are proposed.

**In order to improve the reliability of the depth map constructed using the supervised methods for improving the depth map obtained by active methods**, a supervised method for improving the accuracy of the depth map based on the Markov random field was proposed by implementing a procedure for assessing the mismatch between the edges of the original color image and the resulting depth map with hard and soft problem solving.

As a basic method, we use a method for improving the accuracy of the depth map based on a Markov random field, to which we add a procedure for estimating the edge misalignment between the color image and the depth map obtained by the active method.

Unmatched borders can be divided into two types: the first is made worse by brute force interpolation but can be corrected by pointing to high quality color boundaries, and the second is edge inconsistency. When using a brute force decision, these two cases are classified. Based on the assumption that the number of errors should be less than the capabilities of error-corrected editing, it is assumed that a deep edge is worsened by brute-force interpolation when the offset between it and the nearest colored edge is less than a threshold value. Otherwise, it is interpreted as a true edge mismatch.

Although the tight decision approach produces favorable results for small sampling rates, the performance of the method suffers significantly for large sampling rates. The reason is that the hard-sampling edge disparity estimation cannot work well when the quality of the roughly interpolated depth map is very low. Thus, we propose a mild resolution approach that more precisely measures the degree of mismatch that occurs between depth edges and the corresponding color edges in quantitative terms.

Experimental studies of the proposed method on different data sets have shown an increase in the accuracy of the final depth map by 4.3% - 22.19% for interposition coefficients 2 - 8 for the hard decision-oriented method, and an increase in the accuracy of the final depth map by 5.4% - 7.7% for interposition coefficients 8 - 16 for the softer solution.

**The fourth point of scientific novelty was formulated:** a supervised method for enhancing the precision of a depth map based on a Markov random field *has been improved* by implementing a procedure for evaluating the mismatch between the boundaries of the original color image and the resulting depth map with rigid and flexible problem solving. Using the proposed method, we improved the accuracy of the depth map.

**To improve the precision of a depth map constructed using supervised methods for enhancing a depth map with a coarse sampling ratio obtained by active methods,** a supervised method for improving the prediction accuracy of a low-quality depth map based on a Markov random field was proposed by implementing an boundary misalignment assessment procedure with gentle decision making and representing the

local structure of the depth map using a sparsity forest. Using the proposed method, we can improve the accuracy of the depth map with a large sampling rate.

The proposed method uses a minimal skeletal forest as a factor filter, which is used to automatically remove two unrelated pixels that are close to each other in the spatial domain, making the distance between trees an edge-aware metric. To prevent the «edge bleeding» problem, the proposed method builds a minimal spanning tree for each superpixel in the color image produced by the excessive segmentation.

Experimental studies of the proposed method on different datasets showed an increase in the accuracy of the final depth map by 2.8% – 11.87% for implication coefficients 8 – 16 compared to the most effective analogous methods, including the previously proposed method with mild problem solving..

**The fifth point of scientific novelty was formulated:** a controllable method for increasing the accuracy of a low-quality depth map using a Markov stochastic field *was improved* by introducing an boundary misalignment assessment procedure with mild problem solving and representing the regional structure of the depth map using a minimal skeletal forest. Using the proposed method, we improved the accuracy of the depth map with a large sampling rate

In the **fourth chapter** of the thesis, the proposed methods for constructing a depth map as part of the stereo vision subsystem of a mobile robotic platform were fully tested. Based on the developed experimental stand, which includes a single-board computer, a camera module, and a ToF sensor, it was possible to implement a full cycle of depth reconstruction - from processing input images to obtaining a depth map in real time. This made it possible to test the effectiveness of the developed methods not only in a theoretical or laboratory context, but also in conditions close to practical application in autonomous navigation tasks.

The testing was carried out on data obtained from both standard sets and real sensor equipment, which made it possible to evaluate the performance of the algorithms in terms of accuracy, stability, power consumption, and computational complexity. All methods were implemented taking into account the limitations of embedded systems, in particular, the use of hardware resources, energy efficiency, and real-time processing requirements.

In general, the results of the testing indicate the practical applicability of the proposed methods for implementation in real-time systems. The conclusions drawn from experimental measurements and comparisons with basic approaches confirm the feasibility of their use in vision systems.

*Keywords:* depth maps, technical vision systems, computer vision, image processing, spatial analysis, 3D reconstruction, stereo vision, depth sensors, depth estimation.

### **List of publications of the applicant on the topic of the dissertation**

*Scientific works in which the main scientific results of the dissertation are published:*

1. Arsirii, O.O., Yadrova, M.V., Kondratyev, S.B., & Stelmakh, D.E. Development of the intelligent software and hardware subsystem for capturing an object by robot manipulator. Herald of Advanced Information Technology. 2020; 3(2) 42–51. doi.org/10.15276/hait.02.2020.4. *The publication is included in the list of scientific professional publications of Ukraine, category "B".*

<https://hait.od.ua/index.php/journal/article/view/35>

2. Kondratiev S.B., Kostenko V.L., Yadrova M.V. Contour method for object positioning in mobile computer vision systems. Bulletin of the National Technical University "KhPI". 2021; 2(8): 62–69. doi.org/10.20998/2413-4295.2021.02.09.

*The publication is included in the list of scientific professional publications of Ukraine, category "B".*

[https://www.researchgate.net/publication/352752306\\_CONTOUR\\_METHOD\\_FOR\\_POSITIONING\\_OBJECTS\\_IN\\_MOBILE\\_COMPUTER\\_VISION\\_SYSTEMS](https://www.researchgate.net/publication/352752306_CONTOUR_METHOD_FOR_POSITIONING_OBJECTS_IN_MOBILE_COMPUTER_VISION_SYSTEMS)

3. Antoshchuk, S.G., Kondratyev, S.B., Shcherbakova, G.Y., & Hodovychenko, M.A. Depth map generation for mobile navigation systems based on objects localization in images. Herald of Advanced Information Technology. 2022. 5(1). C. 11–18. doi.org/10.15276/hait.05.2022.1. *The publication is included in the list of scientific professional publications of Ukraine, category "B".*

<https://hait.od.ua/index.php/journal/article/view/128>

4. Kondratyev, S.B., Antoshchuk, S.G. & Hodovychenko, M.A. Methods for refining the depth map obtained from depth sensors. *Applied Aspects of Information Technology*. 2024. № 7(4). C. 336–347. doi.org/10.15276/hait.05.2022.1. *The publication is included in the list of scientific professional publications of Ukraine, category "B"*.

<https://aait.od.ua/index.php/journal/article/view/261>

5. Kondratiev S.B. A local method for constructing a depth map based on an adaptive matching scheme and affine invariant features. *Science and technology today*. 2024; 12(40): 1245–1258. doi.org/10.52058/2786-6025-2024-12(40)-1245-1258. *The publication is included in the list of scientific professional publications of Ukraine, category "B"*.

<http://perspectives.pp.ua/index.php/nts/article/view/17042>

6. Kondratyev, S.B., Antoshchuk, S.G., & Hodovychenko, M.A. Robust depth map refining using color image. *Herald of Advanced Information Technology*. 2024; № 7(4): P. 361–370. doi.org/10.15276/hait.07.2024.25. *The publication is included in the list of scientific professional publications of Ukraine, category "B"*.

<https://hait.od.ua/index.php/journal/article/view/232>

*Scientific papers certifying the approbation of the dissertation materials:*

7. Shcherbakova, G., Antoshchuk, S., Koshutina, D., Sakhno, K., Kondratiev, S. (2025). Adaptive Clustering for Distribution Parameter Estimation in Technical Diagnostics. In: Dovgyi, S., Siemens, E., Globa, L., Kopiika, O., Stryzhak, O. (eds) *Applied Innovations in Information and Communication Technology. ICAIIT 2024. Lecture Notes in Networks and Systems*, vol 1338. Springer, Cham. doi.org/10.1007/978-3-031-89296-7\_19. *The publication is indexed in the Scopus database.*

[https://icaiit.org/proceedings/12th\\_ICAIIT\\_1/2\\_8\\_ICAIIT\\_2024\\_Part\\_2\\_paper\\_6.pdf](https://icaiit.org/proceedings/12th_ICAIIT_1/2_8_ICAIIT_2024_Part_2_paper_6.pdf)

8. Antoshchuk S., Shcherbakova G., Kondratyev S., Koshutina, D., Usov O. Wavelet transform based optimization method for Three-Dimensional computer vision //

Proceedings of the 12th International Conference Information Control Systems & Technologies (ICST 2024), Odesa, Ukraine, September 23–25, 2024, pp. 471–482. *The publication is indexed in the Scopus database.*

<https://ceur-ws.org/Vol-3790/paper41.pdf>

9. Kondratiev S., Koshutina D., Liubomska O., Baskov I. Research on Energy Efficiency of Wi-Fi IoT Systems on Renesas DA16200 Platform // Proceedings of International Conference on Applied Innovation in IT, Koethen, Germany, March 7, 2024, 12(1), pp. 143–148. *The publication is indexed in the Scopus database.*

[https://icaiit.org/proceedings/12th\\_ICAIIT\\_1/2\\_11\\_ICAIIT\\_2024\\_Part\\_2\\_paper\\_9.pdf](https://icaiit.org/proceedings/12th_ICAIIT_1/2_11_ICAIIT_2024_Part_2_paper_9.pdf)

10. Antoshchuk S.G., Shcherbakova G.Yu., Kondratiev S.V., Koshutina D.V. An approximate method for object depth estimation based on wavelet transform. Informatics. Culture. Technology. 2024; 1(1): 197–201.

<https://ics60.aait.od.ua/zbirnik2024.pdf>

11. Svitlana Antoshchuk, Halyna Shcherbakova, Serhii Kondratiev, Daria Koshutina, Oleksandr Usov. Optimization method based on wavelet transform for approximate depth estimation methods // Information control systems and technologies (ICST-Odesa-2024): Proceedings of the XII International Scientific and Practical Conference, September 23-25, 2024, Odesa (Ukraine). ed: Helvetica Publishing House, 2024. p.123–127.

<http://icst-conf.com/2024.pdf>

12. Antoshchuk S.G., Arsirii O.O., Kondratiev S.B. et al. Hybrid representation models and methods of image processing in wavelet transform space. 2020. 160 p.

<https://discovery.kpi.ua/Record/000634213>

13. Antoshchuk S.G., Arsirii O.O., Kondratiev S.B. et al. Image analysis and recognition in the wavelet transform space. 2022. 126 p.

<http://www.irbis-nbuv.gov.ua/publ/REF-0000823682>

## ЗМІСТ

Вступ.....	27
1 Аналіз проблеми побудови карти глибин для систем технічного зору .....	33
1.1 Визначення систем технічного зору .....	33
1.1.1 Класифікація систем технічного зору .....	35
1.1.2 Використання карт глибини в системах технічного зору.....	37
1.2 Методи побудови карти глибин .....	38
1.3 Теоретичні відомості та супутні наукові роботи.....	44
1.3.1 Пасивне визначення глибини за допомогою стерео співставлення .....	44
1.3.2 Базові операції пасивної оцінки глибини .....	46
1.3.3 Теоретичні основи активного визначення глибини за допомогою сенсорів .	50
1.3.4 Локальні методи та фільтри.....	52
1.3.5 Глобальна оптимізація та графічні моделі .....	56
1.3.6 Надроздільна реконструкція карти глибини за допомогою методів на базі навчання.....	62
1.3.7 Використання вейвлет-перетворення при побудові карти глибин .....	65
1.4 Шляхи удосконалення побудови карти глибин для систем технічного зору .	667
1.5 Висновки до першого розділу .....	70
2 Підвищення ефективності побудови карти глибини за допомогою пасивних методів .....	72
2.1 Метод побудови карти глибини з використанням адаптивної процедури зіставлення та афінно-інваріантних ознак .....	72
2.1.1 Афінно-інваріантні ознаки.....	74
2.1.2 Початкове оцінювання глибини.....	76
2.1.3 Уточнення карти глибини.....	82
2.1.4 Експериментальне дослідження запропонованого методу .....	84
2.2 Покращення часової узгодженості для послідовностей карт глибини .....	87
2.2.1 Принцип роботи запропонованого методу .....	88
2.2.2 Експериментальне дослідження роботи методу.....	92

2.3	Метод наближеного стереозіставлення на основі алгоритму StereoBM з використанням вейвлет-перетворення Гаара.....	95
2.3.1	Формальна постановка задачі та суть методу.....	96
2.3.2	Експериментальне дослідження запропонованого методу .....	98
2.3.3	Обговорення результатів дослідження.....	103
2.4	Висновки до другого розділу .....	104
3	Методи удосконалення карт глибини, отриманих за допомогою активних методів.....	107
3.1	Перешкоди в процесі керованого удосконалення карт глибини .....	109
3.2	Керований метод покращення карти глибини з жорсткою оцінкою неузгодженості країв .....	110
3.2.1	Визначення функції енергії Марківського випадкового поля.....	111
3.2.2	Експериментальне дослідження запропонованого методу .....	114
3.3	Керований метод покращення карти глибини з м'якою оцінкою неузгодженості країв .....	115
3.3.1	Модифікована функція енергії Марківського випадкового поля .....	115
3.3.2	Вимірювання неузгодженості країв з м'яким прийняттям рішень .....	116
3.3.3	Включення вимірювань неузгодженості країв у Марківське випадкове поле .....	125
3.3.4	Дослідження складності методу.....	127
3.3.5	Експериментальне дослідження запропонованого методу .....	127
3.3.5.1	Експеримент на наборі даних із синтетичними деградаціями.....	129
3.3.5.2	Експерименти з удосконалення карти глибини з використанням реальних наборів даних.....	134
3.3.5.3	Експерименти з вирішення завдань надроздільної реконструкції карти глибини та побудови карти глибини.....	137
3.4	Удосконалення карти глибини за допомогою використання мінімального кістякового лісу та оцінки неузгодженості країв з м'яким прийняттям рішень ...	138
3.4.1	Модифікація функції енергії мінімального кістякового лісу .....	139
3.4.2	Принцип роботи запропонованого методу .....	140

3.4.2.1 Побудова мінімального кістякового лісу .....	141
3.4.2.2 Вбудовування вимірювання неузгодженості країв у мінімальний кістяковий ліс .....	144
3.4.2.3 Адаптація пропускної здатності для обчислення спорідненості в Марківському випадковому полі .....	145
3.4.3 Експериментальне дослідження запропонованого методу .....	148
3.4.3.1 Дослідження впливу значень параметрів методу на його ефективність...	148
3.4.3.2 Дослідження роботи методу в процесі надроздільної реконструкції карти глибини .....	149
3.4.3.3 Дослідження роботи методу в процесі добудови карти глибини .....	154
3.4.3.4 Дослідження середнього часу роботи методу .....	157
3.5 Висновки до третього розділу .....	158
4 Розробка інструментальних засобів, що реалізують запропоновані рішення та апробація отриманих теоретичних результатів .....	161
4.1 Прикладні аспекти побудови карт глибини на базі удосконалених пасивних методів для автономних робототехнічних систем .....	161
4.1.1 Експериментальний стенд .....	162
4.1.2 Апробація пасивного локального методу на базі адаптивного зіставлення та афінно-інваріантних ознак .....	164
4.1.3 Апробація методу покращення часової узгодженості карт глибини.....	166
4.1.4 Апробація модифікації методу StereoBM з використанням наближеного вейвлет-перетворення Гаара.....	167
4.1.5 Апробація запропонованих рішень на практиці.....	169
4.1.6 Обговорення результатів апробації .....	170
4.2 Прикладні аспекти удосконалення карт глибини, отриманих активними методами.....	171
4.2.1 Експериментальний стенд .....	172
4.2.2 Аналіз кількісних результатів апробації .....	173
4.2.3 Аналіз якісних результатів апробації .....	178
4.2.4 Аналіз енергоспоживання та обчислювальної складності .....	180

4.2.5 Обговорення результатів апробації .....	184
4.3 Висновки до четвертого розділу .....	185
Висновки.....	187
Список використаних джерел.....	190
Додаток А Список публікацій здобувача.....	205
Додаток Б Довідка про впровадження результатів дисертаційної роботи у діяльність НВП «Каре» .....	208
Додаток В Акт впровадження результатів у дисертаційного дослідження у науково-практичну діяльність ТОВ НВО «Діскрет» .....	209
Додаток Г Довідка про використання результатів дисертаційної роботи у науково-дослідницькій діяльності Національного університету «Одеська Політехніка».....	210
Додаток Д Довідка про використання результатів дисертаційної роботи у навчальному процесі Національного університету «Одеська Політехніка» .....	211
Додаток Е Фрагменти вихідного коду програмних інструментальних засобів ...	212

## ВСТУП

**Актуальність теми.** Упродовж останніх десятиліть технічний прогрес у сфері комп'ютерного зору суттєво змінив підходи до автоматизації багатьох процесів у промисловості, медицині, транспорті, безпеці, сільському господарстві та побуті. Розширення функціональності систем технічного зору стало можливим завдяки переходу від аналізу двовимірних зображень до тривимірного просторового сприйняття навколишнього середовища. Одним із ключових елементів такого сприйняття є карта глибини – структуроване представлення сцени, у якому кожному пікселю відповідає значення відстані до відповідної точки об'єкта.

Карта глибини є базовим джерелом просторової інформації, необхідної для низки завдань, зокрема: оцінки положення об'єктів, визначення їхніх розмірів і форм, сегментації сцен, реконструкції тривимірних моделей, оцінки відстаней у середовищі, реалізації навігації автономних агентів, взаємодії з об'єктами в режимі реального часу тощо. Отже, точність, деталізованість і достовірність карти глибини визначають якість роботи всієї системи технічного зору.

У науковій і прикладній літературі описано чимало методів побудови карти глибини: стереозір, структуроване світло, час-прольоту (ToF), методи на основі активного сканування, а також численні алгоритми, що застосовують машинне навчання та глибокі нейронні мережі. Кожен із цих підходів має свої переваги й недоліки – від високої точності до чутливості до шумів, від високої обчислювальної складності до недостатньої узагальнюваності на різні типи сцен. Особливої складності набуває побудова карти глибини у разі слабо текстурованих поверхонь, змін освітлення, наявності прозорих або дзеркальних об'єктів, а також у динамічному середовищі з обмеженими ресурсами.

Таким чином, існує протиріччя, яке полягає в тому, що з одного боку, сучасні системи технічного зору потребують високоякісних карт глибини з високою просторовою та глибинною роздільною здатністю для точного аналізу сцени; з іншого боку, існуючі методи побудови глибинних карт не

забезпечують необхідного рівня точності та стійкості в умовах реального середовища, особливо при обмежених обчислювальних ресурсах або складних зорових сценах.

Тому актуальними є дослідження, що направлені на усунення цього протиріччя та присвячені вирішенню важливої науково-практичної задачі удосконалення побудови карти глибин для систем технічного зору.

**Зв'язок роботи з науковими програмами, планами, темами.** Дисертаційну роботу виконано у відповідності до пріоритетних тематичних напрямів наукових досліджень і науково-технічних розробок на період до 31 грудня року, наступного після припинення або скасування воєнного стану в Україні (Постанова Кабінету Міністрів України від 30.04.2024 р. № 476), згідно пп. 1.2.1.1, 1.2.1.4, 1.2.4.6 і 1.2.1.7 «Основних наукових напрямів та найважливіших проблем фундаментальних досліджень у галузі природничих, технічних, суспільних і гуманітарних наук Національної академії наук України на 2019–2023 роки» (постанова Президії НАН України від 30.01.2019 р. № 30) та також у відповідності до планів науково-дослідних робіт Національного університету «Одеська політехніка» при виконанні держбюджетних тем: НДР роботи № 719-61/161 – «Інтелектуальна підтримка прийняття рішень при проєктуванні кісткових замінників для лікування воєнних травм» (номер держ. реєстрації 0124U000388); НДР № 716-145 «Інноваційна інтелектуальна технологія аналізу та синтезу енергоефективних гідроаеродинамічних елементів та систем» (номер держ. реєстрації 0121U111720). Дисертант приймав участь у виконанні вказаних тем як співвиконавець (акт впровадження від 25.09.2024 р.) (Додаток Г).

**Метою дослідження** є підвищення ефективності побудови карти глибин для систем технічного зору шляхом удосконалення пасивних та активних методів побудови карт глибин.

Для досягнення мети дослідження необхідно вирішити такі завдання:

– провести аналіз основних проблем, обмежень та викликів, які виникають в процесі побудови карт глибин в системах технічного зору;

– удосконалити пасивний локальний метод побудови карти глибини за рахунок використання процедури адаптивного зіставлення та афінно-інваріантних ознак;

– удосконалити метод покращення часової узгодженості для послідовностей карт глибини за рахунок впровадження показника надійності глибини та використання динамічних атрибутів регіонів;

– удосконалити керований метод покращення точності карти глибини за рахунок впровадження процедури оцінки неузгодженості країв вихідного кольорового зображення;

– удосконалити керований метод покращення точності карти глибини низької якості на базі Марківського випадкового поля;

– удосконалити метод локального стереозіставлення StereoVM для систем технічного зору;

– розробити інструментальні засоби, що реалізують запропоновані рішення та провести експериментальні дослідження, апробацію і впровадження отриманих теоретичних результатів.

**Об'єктом дослідження** є процеси формування карт глибини для систем технічного зору.

**Предметом дослідження** є методи підвищення ефективності побудови карти глибин для систем технічного зору.

**Методи дослідження.** Для вирішення поставлених задач використано методи комп'ютерного зору та математичний апарат стохастичного моделювання, зокрема Марківські випадкові поля для формалізації покращення карти глибини як задачі енергетичної оптимізації. Застосовано афінно-інваріантні локальні ознаки для побудови адаптивного зіставлення в пасивних стереосистемах. Для покращення часової узгодженості використано динамічні атрибути регіонів та показник надійності глибини. Методи аналізу локальної структури сцени реалізовано з використанням мінімального кістякового лісу.

**Наукова новизна отриманих результатів.** Основний науковий результат полягає у розв'язанні важливої науково-практичної задачі підвищення ефективності побудови карт глибини в системах технічного зору за рахунок удосконалення методів адаптивного локального зіставлення з використанням афінно-інваріантних ознак, покращення часової узгодженості послідовностей карт глибини з урахуванням динамічних атрибутів сцен та впровадження керованих процедур підвищення точності на основі Марківських випадкових полів з урахуванням узгодженості країв і локальної структури зображення.

У рамках виконаних досліджень отримано такі наукові результати:

– *удосконалено* пасивний локальний метод побудови карти глибини за рахунок використання процедури адаптивного зіставлення та афінно-інваріантних ознак, що дозволило підвищити швидкість побудови карти глибини;

– *удосконалено* метод покращення часової узгодженості для послідовностей карт глибини за рахунок впровадження показника надійності глибини та використання динамічних атрибутів регіонів, що дозволило покращити часову узгодженість та покращити ефективність кодування карти глибини;

– *набув подальший розвиток* метод локального стереозіставлення StereoBM шляхом інтеграції наближеного вейвлет-перетворення Гаара для попереднього виділення контурних ознак та локалізації екстремумів, що дозволило зменшити обчислювальну складність процесу побудови карти глибини, зберігаючи при цьому точність у слабо текстурованих зонах;

– *удосконалено* керований метод покращення точності карти глибини на базі Марківського випадкового поля за рахунок впровадження процедури оцінки неузгодженості країв вихідного кольорового зображення та отриманої карти глибини з жорстким та м'яким прийняттям рішень, що дозволило покращити точність карти глибини;

– *удосконалено* керований метод покращення точності карти глибини низької якості на базі Марківського випадкового поля за рахунок впровадження процедури оцінки неузгодженості країв з м'яким прийняттям рішень та представленням

локальної структури карти глибин з використанням мінімального кістякового лісу, що дозволило покращити точність карти глибини з великим коефіцієнтом дискретизації.

**Практичне значення одержаних результатів.** Розроблені в дисертаційній роботі моделі та методи побудови карти глибини з використанням адаптивного локального зіставлення, афінно-інваріантних ознак, динамічних атрибутів регіонів та Марківських випадкових полів були використані для підвищення точності, швидкодії та часової узгодженості карт глибини в системах технічного зору.

Розроблене програмне забезпечення покращення та аналізу карт глибини дозволяють адаптувати запропоновані рішення до різних задач комп'ютерного зору, зокрема для відеоспостереження, 3D-реконструкції та систем автономної навігації.

Розроблені в роботі методи та інструментальні засоби отримали впровадження у діяльності науково-виробничого підприємства «КАРЕ» (довідка про впровадження результатів від 16 грудня 2024 року), товариства з обмеженою відповідальністю «Діскрет» (акт впровадження результатів від 27 листопада 2024 року) та знайшли відображення у навчальному (довідка про впровадження результатів від 4 грудня 2024 року) та науково-дослідницькому процесі (довідка про впровадження результатів № 35) Національного університету «Одеська політехніка».

**Особистий внесок здобувача.** Основні наукові положення, висновки і рекомендації, що містяться в дисертації та виносяться на захист, отримано здобувачем особисто в період з 2020 по 2025 рік. У статті [1] внесок здобувача полягає в проведенні дослідження стосовно використання карт глибини в мобільних автономних системах, в статті [2] – проаналізовані варіанти підвищення ефективності побудови карт глибини за допомогою пасивних методах в мобільних системах технічного зору, у статті [3] – проведено дослідження ефективності побудови карт глибини в мобільних навігаційних системах за допомогою локалізації об'єктів, у статті [4] – запропонований керований метод покращення точності карти глибини за рахунок впровадження процедури оцінки

неузгодженості країв вихідного кольорового зображення та отриманої карти глибини, у статті [5] – запропонований пасивний локальний метод побудови карти глибини за рахунок використання процедури адаптивного зіставлення та афінно-інваріантних ознак та метод покращення часової узгодженості для послідовностей карт глибини за рахунок впровадження показника надійності глибини та використання динамічних атрибутів регіонів, у статті [6] – запропонований керований метод покращення точності карти глибини на базі Марківського випадкового поля за рахунок впровадження процедури оцінки неузгодженості країв та використання мінімального кістякового лісу, у роботі [8] – дослідження можливості побудови карт глибини на віддаленому сервері для систем інтернету речей, у роботах [7 - 13] – дослідження побудови карт глибини за допомогою вейвлет-перетворення.

**Апробація результатів дисертації.** Основні положення і практичні результати дисертаційної роботи доповідалися та одержали схвалення на таких конференціях:

– міжнародна конференція «International Conference on Applied Innovations in IT 2024 (ICAIT 2024)» (м. Кетен, Німеччина, 2024 рік);

– X Міжнародної науково-практичної конференції Інформатика. Культура. Техніка. 2024 25–27 вересня 2024 р. м. Одеса (Україна);

– XII Міжнародна науково-практична конференція «Інформаційні управляючі системи та технології (ІУСТ - Одеса - 2024)» (м. Одеса, 2024 рік).

**Публікації.** Основні результати дисертаційної роботи викладено в 13 публікаціях з них: 6 статей у наукових фахових виданнях України категорії «Б», 7 публікацій у працях і матеріалах міжнародних наукових конференцій (матеріали трьох з яких проіндексовано у наукометричній базі Scopus), та монографіях. 1 стаття надрукована без співавторів.

**Структура і обсяг роботи.** Дисертація складається з вступу, чотирьох розділів, висновків, списку використаних джерел і додатків. Повний обсяг дисертації складає 224 сторінок, в тому числі: 163 сторінки основного тексту, 43 рисунки і 30 таблиць, список використаних джерел зі 116 найменувань і 6 додатків.

# 1 АНАЛІЗ ПРОБЛЕМИ ПОБУДОВИ КАРТИ ГЛИБИН ДЛЯ СИСТЕМ ТЕХНІЧНОГО ЗОРУ

## 1.1 Визначення систем технічного зору

Системи технічного зору (СТЗ) – це сучасна технологія, яка розвивається на стику обчислювальної техніки, штучного інтелекту, робототехніки та оптики. Їхня основна мета – імітація людського зору для автоматизації процесів аналізу візуальної інформації. Ці системи здатні захоплювати, обробляти та інтерпретувати візуальні дані, допомагаючи вирішувати складні завдання з високою точністю та швидкістю. Вони використовуються у багатьох галузях, таких як промисловість, медицина, транспорт, безпека та сільське господарство, і суттєво змінюють підходи до виконання завдань (рис. 1.1) [1].

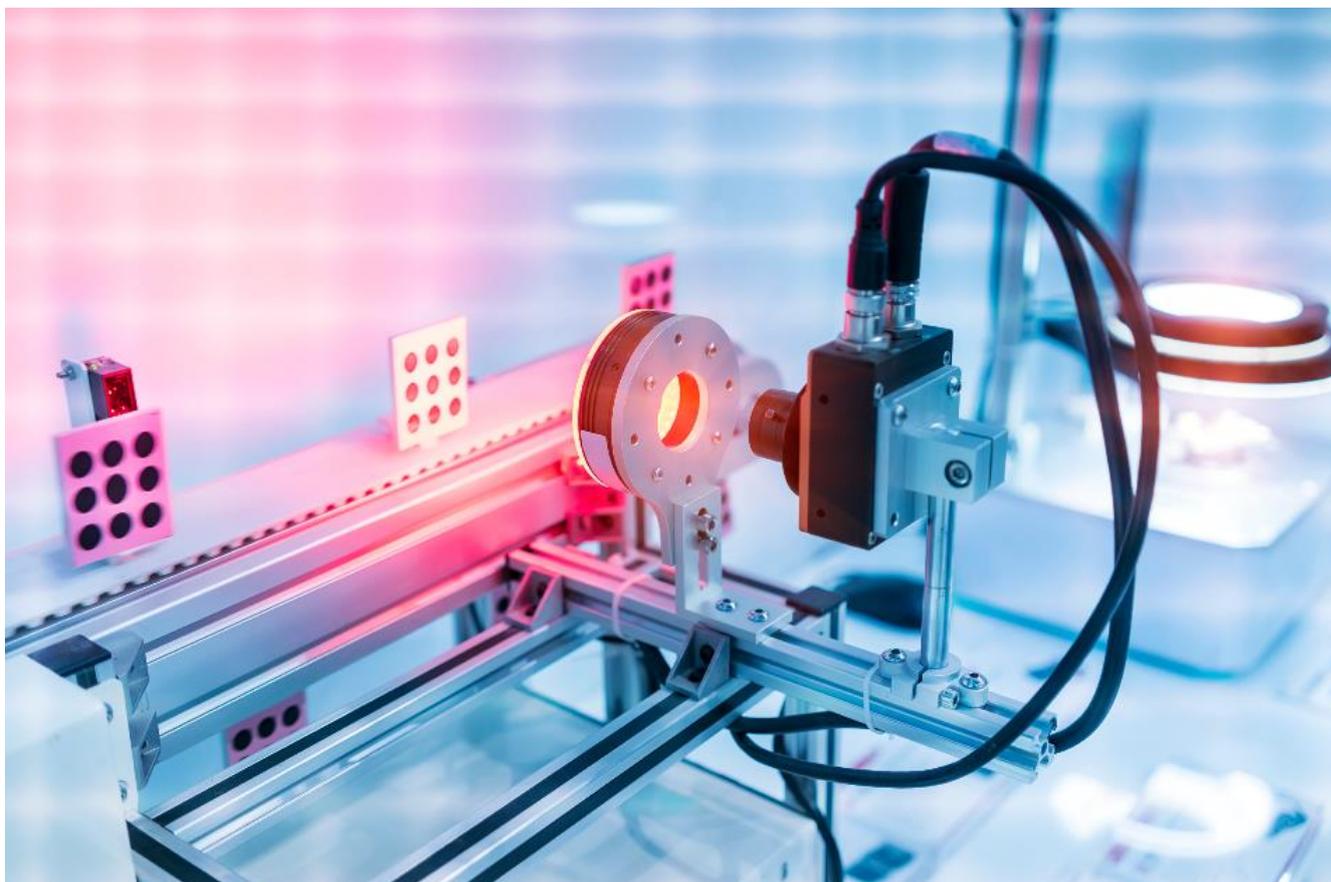


Рисунок 1.1 – Приклад системи технічного зору

Системи технічного зору можна визначити як програмно-апаратні комплекси, які отримують зображення чи відео за допомогою камер та інших сенсорів, обробляють ці дані й аналізують їх для отримання необхідної інформації. У порівнянні з традиційними сенсорами, які вимірюють конкретні фізичні параметри, СТЗ здатні виконувати більш складні завдання, такі як розпізнавання об'єктів, оцінка їхньої форми, кольору, розмірів або аналіз руху. Це робить їх надзвичайно універсальними та ефективними у вирішенні різноманітних задач.

Розвиток систем технічного зору почався ще в 1960-х роках, коли з'явилися перші цифрові камери й алгоритми для аналізу зображень. У 1970-х роках ці технології почали застосовуватися в промисловості, зокрема для контролю якості продукції. Протягом 1980-х років із розвитком обчислювальної техніки з'явилися перші інтегровані рішення для автоматизації виробничих процесів. У 2000-х роках, завдяки прогресу в машинному навчанні та штучному інтелекті, СТЗ отримали можливість автоматично розпізнавати образи та класифікувати об'єкти. Сьогодні ці системи поєднуються з іншими технологіями, такими як роботи, безпілотні транспортні засоби та хмарні обчислення, що робить їх ще більш потужними та гнучкими [2].

Системи технічного зору знаходять застосування у найрізноманітніших сферах. У промисловості вони використовуються для контролю якості продукції, сортування матеріалів і вимірювання параметрів деталей. У медицині СТЗ допомагають аналізувати рентгенівські знімки, МРТ та інші медичні зображення, а також забезпечують роботу хірургічних роботів. У транспорті системи технічного зору є невід'ємною частиною автономних автомобілів і систем допомоги водіям, таких як виявлення перешкод або дорожніх знаків. У сфері безпеки вони забезпечують розпізнавання облич, відеоспостереження та аналіз поведінки. У сільському господарстві ці системи використовуються для моніторингу стану врожаю, автоматизації збору й аналізу продукції.

Системи технічного зору є надзвичайно важливими, оскільки вони забезпечують автоматизацію складних процесів, які раніше виконувалися людиною. Їхня здатність працювати з великими обсягами візуальних даних у

реальному часі відкриває нові можливості для підвищення ефективності, точності та швидкості багатьох процесів. Ці технології продовжують розвиватися, інтегруючись із новітніми досягненнями науки, і, безсумнівно, матимуть ще більший вплив на наше майбутнє [3].

### 1.1.1 Класифікація систем технічного зору

Системи технічного зору можна класифікувати за різними критеріями, що лежать в її основі, та характеризують їхню функціональність, архітектуру, тип обробки даних, сферу застосування тощо. Розглянемо основні види класифікації, які дозволяють зрозуміти особливості певних типів систем і їхню роль у конкретних завданнях (табл. 1.1) [4].

*Класифікація СТЗ за призначенням.* Системи технічного зору можуть бути розроблені для виконання конкретних завдань, таких як контроль якості на виробництві, розпізнавання облич, аналіз дорожнього руху чи допомога в роботі автономних транспортних засобів тощо (табл. 1.1).

*Класифікація СТЗ за типом обробки даних.* Системи технічного зору можуть працювати з двовимірними (2D) або тривимірними (3D) даними:

– 2D-системи – обробляють двовимірні зображення або відео, що отримуються з плоских камер. Це достатньо для аналізу текстур, кольорів, форм та контурів;

– 3D-системи – використовують стереокамери, лазерні сканери або глибинні сенсори для створення тривимірних моделей об'єктів і середовища. Це актуально для робототехніки, де потрібна точна інформація про просторове положення об'єктів.

*Класифікація СТЗ за способом інтеграції в процеси:*

– автономні системи – працюють як самостійні пристрої, виконуючи всі етапи від збору зображення до аналізу та прийняття рішення без втручання інших систем;

– інтегровані системи – є частиною більшого комплексу, наприклад, промислової лінії, де результати обробки зображень передаються до центральної системи управління для подальших дій.

Таблиця 1.1 – Класифікація систем технічного зору

Критерій класифікації	Тип системи	Опис
За призначенням	Контроль якості у виробництві	Автоматична перевірка, виявлення дефектів, сортування за параметрами
	Розпізнавання облич	Безпека, ідентифікація, доступ до пристроїв або приміщень
	Навігація автономних транспортних засобів	Орієнтація у просторі, виявлення перешкод, прийняття рішень
За типом обробки даних	2D-системи	Обробка двовимірних зображень, аналіз текстур, кольорів, форм
	3D-системи	Створення тривимірних моделей, точна інформація про просторове положення
За способом інтеграції	Автономні системи	Повна автономія – збір, аналіз і прийняття рішень без зовнішнього втручання
	Інтегровані системи	Є частиною більшої системи, передає результати до централізованого управління
За технологією побудови	Традиційні системи	Класичні алгоритми обробки зображень: фільтрація, сегментація, детекція
	Системи на основі штучного інтелекту	Використання нейронних мереж і ML для адаптації до нових задач
За типом середовища	Контрольовані умови	Стабільні параметри середовища, наприклад, виробництво
	Змінні умови	Змінні умови: вулиці, громадські місця, потреба в адаптації

*Класифікація СТЗ за технологією побудови систем.* Сучасні СТЗ часто базуються на алгоритмах машинного навчання і нейронних мережах, які дозволяють обробляти великі обсяги даних і адаптуватися до нових завдань:

– традиційні системи – використовують алгоритми класичної обробки зображень, такі як фільтрація, сегментація, детекція контурів та інші методи аналізу;

– системи на основі штучного інтелекту – застосовують нейронні мережі та алгоритми машинного навчання, що дозволяє системам адаптуватися до нових завдань і працювати з нетиповими даними або в складних умовах.

*Класифікація СТЗ за типом середовища, у якому працює система:*

– контрольовані умови – системи працюють у середовищі з фіксованими параметрами, наприклад, на виробничій лінії з постійним освітленням та стабільним розташуванням об'єктів;

– змінні умови – використовуються в середовищах, де параметри постійно змінюються (вулиці, громадські місця, відкриті простори), що потребує адаптації до зміни освітлення, погоди чи положення об'єктів.

Таким чином, класифікація систем технічного зору відображає їхню багатогранність та універсальність. Вона допомагає краще зрозуміти, які саме системи слід застосовувати для вирішення певних завдань, враховуючи їхні технічні характеристики, функціональні можливості та обмеження. Кожна класифікація відкриває нові перспективи для розробки та вдосконалення цих технологій, що дозволяє ще більше розширити сфери їхнього застосування.

### **1.1.2 Використання карт глибини в системах технічного зору**

Карти глибини відіграють важливу роль у сучасних системах технічного зору, забезпечуючи тривимірне представлення сцени та дозволяючи здійснювати аналіз просторових характеристик об'єктів.

На відміну від традиційних зображень у видимому спектрі, які містять лише інформацію про інтенсивність або колір, карти глибини відображають відстань до кожної точки сцени відносно сенсора. Це значно розширює можливості комп'ютерного зору, особливо в контексті задач, що потребують просторового розуміння, зокрема виявлення об'єктів, реконструкції сцени, сегментації, навігації автономних систем тощо [5].

Отримання карти глибини можливе за допомогою різноманітних технологій, серед яких слід виокремити стереозір, час прольоту (Time-of-Flight), структуроване

світло та глибинні камери на базі активного інфрачервоного випромінювання. Кожна з технологій має свої переваги та обмеження щодо точності, дальності, стійкості до зовнішніх умов та вартості та особливостей реалізації. Наприклад, стереозір дозволяє отримати карту глибини з двох звичайних камер, проте потребує високої точності в калібруванні та підборі відповідностей. Натомість Time-of-Flight камери надають глибину безпосередньо, але можуть бути чутливими до шуму в умовах яскравого освітлення.

Використання карт глибини суттєво підвищує ефективність алгоритмів сегментації, класифікації об'єктів і побудови тривимірних моделей середовища. Зокрема, в задачах робототехніки глибина дозволяє більш точно оцінювати положення та орієнтацію об'єктів, уникати зіткнень, планувати траєкторії та взаємодіяти з оточенням.

У сфері безпеки та відеоспостереження глибинна інформація сприяє покращенню точності виявлення людей, оцінки пози та руху, навіть в умовах часткового перекриття або низької якості зображення.

Таким чином, карти глибини становлять важливий інструмент у побудові інтелектуальних систем технічного зору, що здатні сприймати світ у трьох вимірах та адаптивно взаємодіяти з динамічним середовищем. Подальший розвиток методів обробки та інтеграції глибинної інформації відкриває нові можливості для підвищення автономності, надійності та інтелектуальності технічних систем [6].

## **1.2 Методи побудови карти глибин**

Глибина представляє геометричну інформацію реальних сцен шляхом знаходження відстаней між пікселями і камерою, яка може трансформувати пікселі зображення з двовимірної системи координат у тривимірний простір (тобто хмару точок). Отримання високоякісних карт глибини є ключовою проблемою в області 3D комп'ютерного зору, яка необхідна в багатьох сучасних додатках, наприклад, в інтерактивній інтерполяції зображення, 3DTV, 3D-моделюванні об'єктів, навігації роботів і 3D-трекінгу, однак ця задача складніша, ніж звичайне отримання

зображень. Загалом, методи отримання карт глибини можна поділити на дві категорії: пасивні та активні.

Пасивні методи можуть генерувати карту глибини з дворакурсних або багаторакурсних кольорових зображень, використовуючи алгоритми стерео співставлення. Протягом десятиліть було запропоновано безліч пасивних методів. Методи стерео співставлення можна розділити на два основні класи: локальні та глобальні.

Локальні методи обчислюють значення глибини для кожного пікселя незалежно, співставляючи лише локальну інформацію. Глобальні методи визначають значення глибини для всіх пікселів одночасно на основі попередньо заданого значення (наприклад, норми повної варіації).

Локальні методи, як правило, швидші за глобальні, але це відбувається за рахунок втрати точності. За останні кілька десятиліть продуктивність стерео співставлення значно покращилася. Однак ці методи все ще страждають від притаманних їм проблем, таких як труднощі співставлення в безтекстурних областях та оклюзії [7], які показані на рисунку 1.2.

Слід відмітити, що алгоритми стереопредставлення можуть бути реалізовані як у вихідному просторі зображення, так і в просторі перетворень (вейвлет-перетворення, Фур'є перетворення).

В реальних додатках важливо збалансувати час виконання та відповідну продуктивність. Крім того, оскільки карти глибини обчислюються кадр за кадром, значення глибини статичного фону в сусідніх кадрах не можуть бути стабільними, що призводить до появи артефактів (наприклад, артефактів миготіння, коли ці карти глибини використовуються для візуалізації віртуального відео). Ця проблема називається часовою неузгодженістю [8].

Активні методи отримання глибини дозволяють отримувати відео з такою ж частотою кадрів, як і кольорові камери, за допомогою сенсорів глибини. Порівняно з пасивними методами, отримання глибини за допомогою активних методів є набагато ефективнішим.



а)  
б)  
Рисунок 1.2 – Ілюстрація оклюзії (а) та безтекстурної ділянки (б)

Зокрема, в безтекстурних областях активні методи здатні досягти більш надійних результатів, ніж пасивні. На сьогоднішній день існує два основних типи датчиків глибини: ToF-сенсори (наприклад, Sony DepthSense IMX556, рис. 1.3) і датчики структурованого світла (наприклад, Orbbec Astra Pro, рис. 1.4).



Рисунок 1.3 – Зображення ToF-сенсору Sony DepthSense IMX556 [9]

Використовуючи датчики ToF, карти глибини обчислюються шляхом вимірювання різниці фаз між випромінюваним і відбитим світлом [10].



Рисунок 1.4 – Зображення сенсору структурованого світла Orbbec Astra Pro [11]

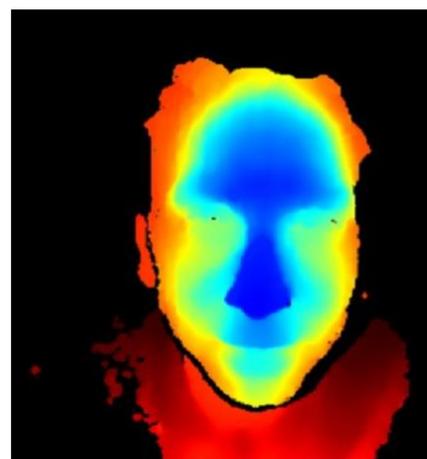
Недоліком використання ToF-сенсорів є те, що отримані карти глибини є зашумленими та мають низьку роздільну здатність (наприклад,  $176 \times 144$  або  $200 \times 200$ , найбільш досконалі сенсори дозволяють отримати карту глибини роздільною здатністю  $640 \times 480$ ).

Використовуючи датчики структурованого світла, інфрачервоне джерело світла проектує точковий візерунок на сцену. Інша інфрачервона камера зі зміщеним фокусом фіксує цей рисунок і оцінює карту глибини. Хоча датчики зі структурованим світлом можуть отримувати карти глибини з вищою роздільною здатністю, якість карт глибини, отриманих за допомогою таких датчиків, не є задовільною. Часто, на таких картах глибини з'являються прогалини (тобто, місця, де інформація про глибину не сприймається).

Ці прогалини можуть бути спричинені оклюзією, слабким відбиттям інфрачервоного світла від поверхонь або навіть тіньовим відбиттям світлових патернів. На рисунку 1.5 наведені приклади карт глибини, отримані ToF-сенсорами та датчиками структурованого світла.



а)



б)

Рисунок 1.5 – Ілюстрація карт глибин, отриманих ToF-сенсором (а) та датчиком структурованої глибини (б)

Загалом, об'єкти темних кольорів, дзеркальні поверхні або дрібнозернисті поверхні, наприклад, людське волосся, важко визначити за допомогою датчиків

глибини [12]. Згідно з наведеним вище аналізом, основними проблемами карт глибини, отриманих за допомогою датчиків глибини, є низька роздільна здатність, зашумлені значення глибини та прогалини.

Такі низькоякісні карти глибини завжди генеруються з зафіксованими високоякісним кольоровим зображенням або зображенням у відтінках сірого (в рамках даного дисертаційного дослідження, кольорове зображення або зображення у відтінках сірого буде називатися кольоровим зображенням).

Тому карту глибини можна покращити, керуючись відповідним кольоровим зображенням завдяки високій кореляції між ними. Такі методи називають керованими методами покращення карти глибини.

Існуючі методи керованого покращення карти глибини можна розділити на різні категорії залежно від того, чи використовуються зовнішні навчальні дані: перші методи здійснюють розв'язання заздалегідь визначеної моделі шляхом оптимізації, тоді як інші навчають моделі на основі зовнішніх даних.

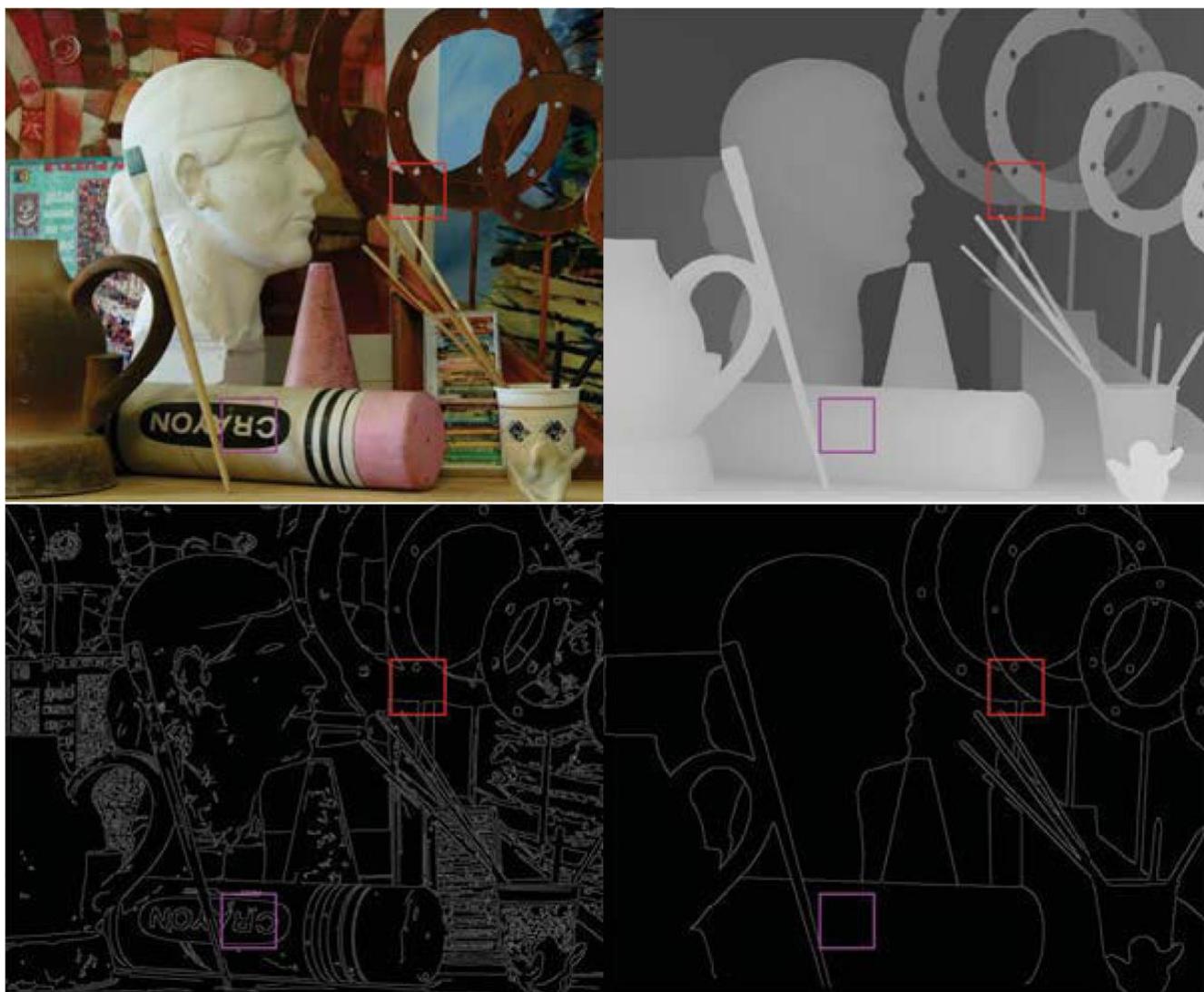
Методи, які не використовують зовнішні дані, часто експлуатують властивість співпадіння країв на карті глибини та на відповідному кольоровому зображенні [13]. Такий підхід дозволяє продемонструвати, що деталізацію і точність більшості країв глибини можна покращити на основі інформації, яку надають відповідні краї кольорового зображення. Однак, це припущення не завжди вірне.

Неправильне орієнтування за супутнім кольоровим зображенням призведе до появи артефактів копіювання текстури та розмиття країв глибини на покращеній карті глибини. Артефакти копіювання текстури виникають через те, що згладжена область глибини відповідає кольоровій області з багатою текстурою.

На противагу цьому, розмиття країв глибини є наслідком того, що гладка кольорова область відповідає області глибини з краями. На рисунку 1.6 наведена ілюстрація ситуації невідповідності країв, яка була описана вище.

Підходи, що використовують зовнішні дані, в першу чергу, засновані на розрідженому представленні [14]. У таких підходах відповідні ділянки зображення, отримані з карти глибини низької роздільної здатності (НРЗ), карти глибини

високої роздільної здатності (ВРЗ) та кольорового зображення, можуть бути представлені за допомогою розрідженого вектора на відповідних вивчених загальних словниках.



(а) кольорове зображення та карта його країв

(б) карта глибини та карта країв

Рисунок 1.6 – Ілюстрація неузгодженості країв: краї у рожевих квадратах є на кольоровому зображенні, але не на карті глибини, краї у червоних квадратах з'являються на глибині на карті глибини, але не на кольоровому зображенні

Припускаючи, що розріджений вектор є спільним для цих зареєстрованих ділянок, статистичні залежності між кольоровими та глибинними ділянками можуть бути вивчені за допомогою спільної схеми розрідженого кодування.

Декілька років тому з'явилося кілька робіт на основі наскрізних згорткових нейронних мереж (CNN) [15]. Порівняно з методами на основі моделей, які явно використовують властивість співпадіння країв на карті глибини і відповідного кольорового зображення, такі методи на основі даних покращують карту глибини низької якості, неявно використовуючи підказки із зареєстрованого кольорового зображення. Через різні статистичні атрибути, артефакти копіювання текстури та розмиття країв карти глибини також можуть виникати в таких методах машинного навчання. Вони значною мірою покладаються на подібність між розподілом навчальних даних і розподілом тестових даних.

### **1.3 Теоретичні відомості та супутні наукові роботи**

#### **1.3.1 Пасивне визначення глибини за допомогою стерео співставлення**

Кожен піксель у певному ракурсі та відповідний йому піксель у іншому ракурсі можуть бути визначені за допомогою стерео співставлення кольорових зображень однієї і тієї ж сцени, отриманих з різних ракурсів. Вектор зсуву між кожною зареєстрованою парою пікселів називається зміщенням.

Пов'язані з цим поняття включають ректифікацію зображення та зв'язок між зміщенням і глибиною. Ректифікація зображення є необхідним кроком попередньої обробки, щоб зменшити розмірність зміщення з двох до одного. Крім того, оскільки результатом стерео співставлення є зміщення, слід звернути увагу на зв'язок між зміщенням та глибиною.

*Ректифікація зображення.* Метою ректифікації зображення є вирівнювання епіполлярних ліній двох зображень по горизонталі. Це можна зробити за допомогою лінійних перетворень зображень з використанням внутрішніх і зовнішніх параметрів камери [16].

Після вирівнювання зображення кандидати розташовуються в тому ж рядку, що й піксель, який потрібно зіставити. Таким чином, розмірність пошуку зменшується з двох до одного. На рисунку 1.7 наведено ілюстрацію процедури

ректифікації зображень. Надалі передбачається, що пари зображень попередньо випрямлені.

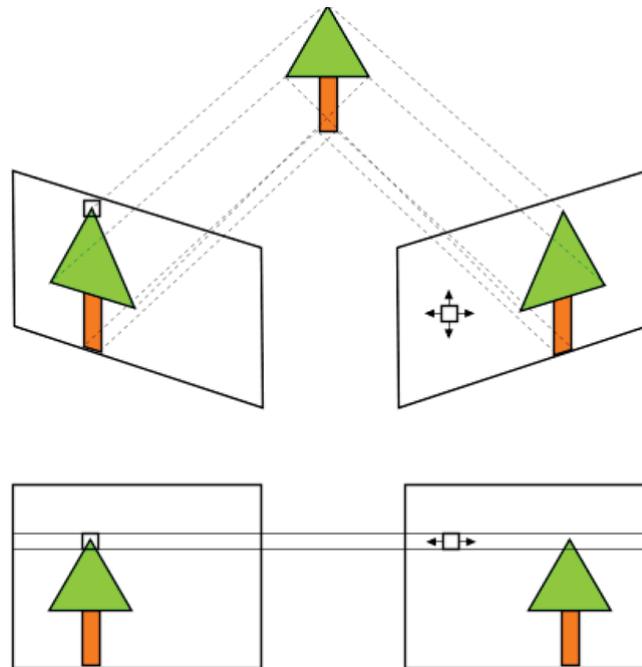


Рисунок 1.7 – Ілюстрація процедури ректифікації зображень

*Зв'язок між зміщенням та глибиною.* Ілюстрація зв'язку між зміщенням і глибиною для ректифікованої пари зображень показана на рисунку 1.8. Виходячи з даної ілюстрації, зміщення  $d$  точки  $M$  можна обчислити за наступною формулою:

$$d = x_L - x_R = \frac{FL_1}{Z_M} + \frac{FL_2}{Z_M} = \frac{FL}{Z_M}, \quad (1.1)$$

де  $F$  і  $L$  – фокусна відстань і базова довжина відповідно,  $x_L$  і  $x_R$  – координати  $x$  проєкцій на дві площини,  $Z_M$  – координата  $z$  точки  $M$ . Це рівняння означає, що після обчислення різниці можна знайти відповідну глибину.

*Обмеження стерео співставлення.* Оскільки щільне оцінювання глибини на основі стерео співставлення є погано сформульованим, необхідно ввести деякі попередні умови для зменшення невизначеності [17].

Подібність інтенсивності. Однією з основних передумов є подібність інтенсивності, яка означає, що відповідні пікселі, які належать до однієї і тієї ж частини об'єкта на декількох зображеннях, мають подібне значення інтенсивності. Цей принцип не спрацьовує, якщо умови освітлення камер відрізняються.

Унікальність. Унікальність означає, що всі пікселі збігаються один до одного. Це не завжди вірно, наприклад, при оклюзії.

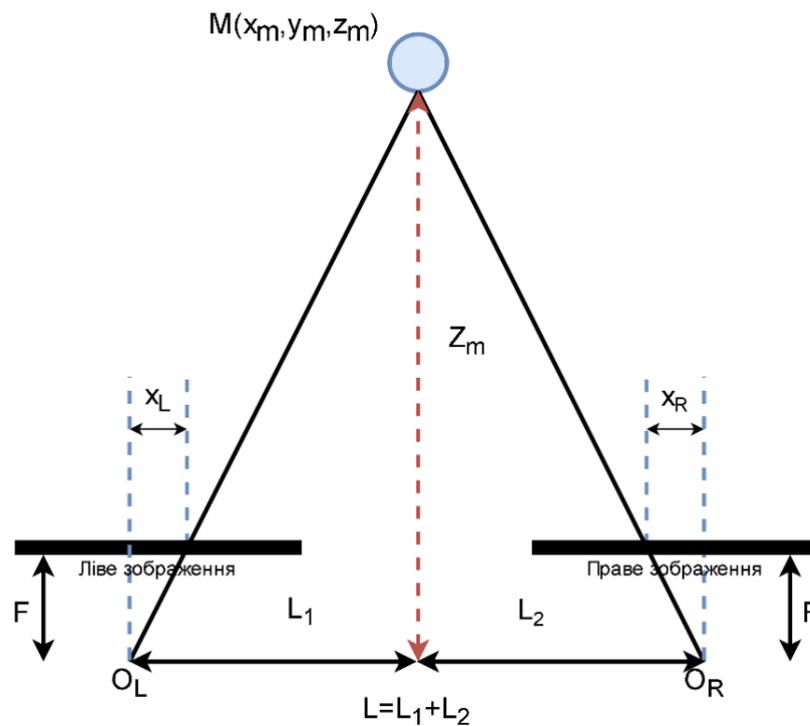


Рисунок 1.8 – Ілюстрація зв'язку між зміщенням та глибиною

Гладкість. Карта глибини є гладкою вздовж поверхні об'єктів. Розриви виникають лише на краях об'єктів.

### 1.3.2 Базові операції пасивної оцінки глибини

До базових операцій пасивної оцінки глибини відносять агрегування вартостей та обчислення зміщення.

Зазвичай, стерео співставлення виконується в межах певного діапазону пошуку зміщення. Якість співставлення, яке відображає піксель поточного ракурсу

на піксель іншого ракурсу, визначеного зміщенням, оцінюється вартістю співставлення, визначеною певною функцією вартості. Обсяг витрат складається з витрат на співставлення всіх пікселів з усіма кандидатами на зміщення у діапазоні пошуку. Ця процедура називається агрегуванням витрат.

Після обчислення такого обсягу витрат можна визначити зміщення локальними або глобальними методами.

*Агрегування витрат.* Для певного зміщення  $d$  у визначеному діапазоні пошуку  $[d_{min}, d_{max}]$  існує вартість зіставлення між еталонним пікселем  $k$  з координатами  $(k_x, k_y)$  та відповідним йому кандидатом  $l$  з координатами  $(l_x = k_x + d, l_y = k_y)$ . Для простоти,  $k$  і  $l = k + d$  будуть визначати пару пікселів, що співпадають. Оскільки співставлення одного пікселя чутливе до шуму, вартість обчислюється між двома локальними вікнами з центрами в  $k$  і  $l$ .

Існує багато функцій вартості, що відповідають різним одиницям виміру оцінки. На основі попередньої схожості за інтенсивністю, запропоновано суму абсолютних різниць інтенсивності (SAD) [18] та суму квадратів різниць інтенсивності (SSD) [19].

Відповідно до ознак в полі градієнта, як доповнення до SAD або SSD представлена сума абсолютних різниць градієнтів (SGRAD), коли попередня оцінка подібності інтенсивності не спрацьовує [20]. Ці функції вартості наведені у відповідних рівняннях:

$$C_{SAD}(k, l) = \sum_{s \in N_k} |Img_L(s) - Img_R(s + d)|, \quad (1.2)$$

$$C_{SSD}(k, l) = \sum_{s \in N_k} (Img_L(s) - Img_R(s + d))^2, \quad (1.3)$$

$$C_{SGRAD}(k, l) = \sum_{s \in N_k} |\nabla_x Img_L(s) - \nabla_x Img_R(s + d)| + |\nabla_y Img_L(s) - \nabla_y Img_R(s + d)|, \quad (1.4)$$

де  $Img_L$  і  $Img_R$  – зображення, зняті з різних ракурсів,  $N_k$  – локальне вікно з центром в  $k$ ,  $\nabla_x$  та  $\nabla_y$  – оператори градієнта,  $d$  – елемент у діапазоні пошуку зміщення  $[d_{min}, d_{max}]$ .

*Обчислення зміщення.* На основі обчисленого обсягу витрат  $C(k, l)$  для кожного пікселя  $k$  і кожного зміщення  $d$  можна визначити оптимальні зміщення для всіх пікселів за допомогою локального або глобального методів. Локальні методи визначають зміщення  $d_k$  для кожного пікселя  $k$  незалежно, щоб знайти найкращу відповідність, яка моделюється рівнянням:

$$d_k = \arg \min_d C(k, l). \quad (1.5)$$

На відміну від локальних методів, глобальні методи завжди явно накладають певну регуляризацію на карту зміщень, наприклад, попереднє згладжування. А оптимальні зміщення  $D^*$  для всіх пікселів визначаються одночасно через глобальну оптимізацію функції енергії. Ця схема має наступний вигляд:

$$D^* = \arg \min_D data(D) + \lambda reg(D), \quad (1.6)$$

$$data(D) = \sum_{d_k \in D} C(k, d_k), \quad (1.7)$$

$$reg(D) = \sum_i \sum_{j \in N_i} V(d_i, d_j), \quad (1.8)$$

де  $N_i$  – околиця пікселя  $i$ ,  $data$  та  $reg$  – член даних та член регуляризації відповідно.  $\lambda$  – коефіцієнт балансу. Загальна ідея члена регуляризації полягає в тому, щоб дозволити  $V$  монотонно зростати з різницею зміщення, щоб накласти штраф на розривний результат, в той час як мати можливість зменшити цей штраф за умови розташування на краях кольорового зображення, щоб зберегти краї глибокого зображення. Такі методи пов'язані з графічними моделями, наприклад, Марковським випадковим полем (Markov Random Field, MRF) [21].

*Уточнення глибини.* Оскільки глибина майже скрізь змінюється безперервно і розриви глибини відбуваються переважно на краях кольорового зображення, значення глибини можна наближено змодельювати як площину в межах сегментів. Цей метод уточнення називається площинною підгонкою. Іншим широко використовуваним методом постобробки є перехресна перевірка, яка перевіряє зміщення за допомогою карт зміщення у різних точках огляду.

Площинна підгонка за допомогою RANSAC. Карта глибин, оброблена площинною підгонкою, завжди має атрибут гладкості. Однак, така підгонка може бути складною через наявність шуму та викидів на початковій знайденій карті глибин. Порівняно з методом найменших квадратів, такі методи, як Random Sample Consensus (RANSAC) [22] є більш стійкими до викидів.

RANSAC чергує два етапи, висування гіпотез і тестування, з визначенням моделі та оціночної функції, яка забезпечує відстань між точкою даних і моделлю. Етап висування гіпотези полягає в тому, щоб підігнати модель, використовуючи мінімальну кількість точок вибірки, необхідну для повного визначення параметрів моделі, які вибираються випадковим чином. На етапі тестування вимірюється відстань між усіма іншими точками даних і моделлю за допомогою оціночної функції. Для класифікації вхідних та вихідних даних визначається порогове значення. Модель з найбільшою кількістю пропусків вважається найбільш репрезентативною. У якості опціонального кроку, модель може бути підігнана з використанням всіх знайдених пропусків.

У випадку площинної підгонки, використовується модель та оціночна функція  $E$  – це рівняння площини та відстань між точкою даних  $P$  з координатами  $(p_x, p_y, p_z, 1)$  до площини  $T$ , представлені нормальним вектором  $(t_x, t_y, t_z, 1)$ , які показані у вигляді наступних рівнянь:

$$S \cdot T = 0, \quad (1.9)$$

$$E = \frac{P \cdot T}{\sqrt{t_x^2 + t_y^2 + t_z^2}}, \quad (1.10)$$

де  $S(s_x, s_y, s_z, 1)$  – довільна точка на площині  $T$ . Кількість ітерацій є параметром, що самоналаштується і залежить від ймовірності  $pg$  знаходження гарної вибірки та максимально допустимої ймовірності  $pb$  вибору лише поганих вибірок. Кількість ітерацій визначається як  $(1 - pg)^T \leq pb$ .

Перехресна перевірка двох карт зміщення. Карта зміщення представляє зміщення від пікселів еталонного зображення до відповідних пікселів цільового зображення. В ідеалі, карта зміщення, на якій еталонне і цільове зображення поміняні місцями, є еквівалентною оригінальній карті. Однак, ненадійні оцінки зміщення для певного набору пікселів на таких картах зміщення завжди суперечать одна одній. Тому для перевірки достовірності зміщення для кожного пікселя можна використовувати карти зміщення для декількох ракурсів [23]. Зокрема, надійність цих карт визначається шляхом порівняння зміщень опорних пікселів зі зміщеннями цільових пікселів на інших ракурсах, що визначається наступною формулою:

$$D_L(x, y) = D_R(x - D_L(x, y), y), \quad (1.11)$$

$$D_R(x, y) = D_L(x + D_R(x, y), y), \quad (1.12)$$

де  $D_L$  та  $D_R$  – карти зміщень для лівого та правого ракурсів відповідно,  $x$  та  $y$  – координати на картах зміщень. Якщо певні пікселі не задовольняють цій перевірці, вони є недостовірними, що можна уточнити за допомогою фільтрів.

### **1.3.3 Теоретичні основи активного визначення глибини за допомогою сенсорів**

Кarti глибини можуть бути отримані за допомогою сенсорів, що є більш стійким до безтекстурних областей порівняно з пасивним зчитуванням глибини. Однак, якість таких карт глибини не задовольняє потреби додатків. Таким чином, ці методи отримання карти глибини також потребують вдосконалення. В рамках даного дисертаційного дослідження, «вдосконалення процесу отримання глибини»

включає в себе надроздільну реконструкцію карти глибини та побудову карти глибини, які були отримані за допомогою ToF-сенсорів та сенсорів структурного світла [24].

*Модель деградації.* Нехай  $Y$  – деградований сигнал або зображення, його модель деградації можна визначити наступним чином:

$$Y = HX + N, \quad (1.13)$$

де  $H$  – матриця деградації, що відповідає зменшенню дискретизації та генерації прогалін,  $X$  – вихідний високоякісний сигнал,  $N$  – адитивний шум.

Така модель дає співвідношення між високоякісною картою глибини та її низькоякісною версією, яка використовується при реконструкції глибини.

*Пов'язані роботи.* Оскільки надроздільну реконструкцію карти глибини та побудову карти глибини можна сформулювати як одну й ту саму обернену задачу, то в даному дисертаційному дослідженні розглядаються загальні методи для цих двох задач. Існуючі методи можна розділити на дві категорії: некеровані методи [25] та керовані методи [26].

Розглянемо роботи, присвячені розробці некерованих методів. Оскільки дисертаційне дослідження зосереджено на керованих методах покращення карти глибини, стисло розглянемо роботи, присвячені некерованим методам.

Автори роботи [27] запропонували метод, який вимагає лише однієї карти глибини для надроздільної реконструкції, використовуючи згладжування на основі локальної самоподібності, але запропонований метод або має труднощі в текстурованих областях, або добре працює лише у випадку невеликого коефіцієнту дискретизації.

Автори роботи [28] запропонували метод надроздільної реконструкції з однією картою глибини за допомогою модифікованого спільного двостороннього фільтра. Такий двосторонній фільтр керується картою країв з високою роздільною здатністю, яка будується з країв карти глибини з низькою роздільною здатністю за допомогою оптимізації Марківського випадкового поля на основі патч-синтезу.

Інший тип некерованого підходу [29] полягає в об'єднанні декількох зміщених карт глибини низької роздільної здатності в одну карту глибини високої роздільної здатності, що не є зручним для реальних застосувань, оскільки геометричне співвідношення між усіма датчиками глибини не може бути легко визначене.

Розглянемо роботи, присвячені розробці керованих методів. Керовані методи мають на меті покращити якість отриманої карти глибини за допомогою зареєстрованого кольорового зображення. Такі керовані методи можна розділити на три категорії: локальні методи [30], глобальні методи [31] та методи на основі навчання [32].

Локальні методи зазвичай уточнюють глибину для кожного пікселя незалежно та інтерполюють поточні недоступні пікселі через середньозважене значення. Адаптивна вага визначається на основі зареєстрованого кольорового зображення.

Глобальні методи виконують покращення шляхом моделювання цієї задачі як задачі оптимізації попередньо визначеної функції енергії. Вони завжди базуються на графічних моделях, які наносять на карту глибини певні попередні орієнтири на основі кольорового зображення, щоб зменшити неоднозначність такої оберненої задачі. Фундаментальним припущенням локальних і глобальних методів є те, що краї карти глибини і краї кольорового зображення узгоджуються у відповідних місцях. Методи на основі навчання використовують зовнішні набори даних, що дозволяє вивчити вказівки і попередні атрибути для конкретного типу сигналу. Це є прямим продовженням покращення окремого зображення. Оскільки існують загальні методи, які стосуються як оцінки глибини, так і покращення глибини, вони детально описані нижче. Ці методи складаються з фільтрів, графічної моделі та глобальної оптимізації.

### **1.3.4 Локальні методи та фільтри**

*Фільтри оптимізації за  $L_2$ -нормою.* Для початку, слід згадати методи двосторонньої фільтрації [33], які можна представити наступним рівнянням:

$$e'_k = \sum_{l \in N_k} W(l, k) \times e_l, \quad (1.14)$$

де  $e_l$  – елемент, що використовується для середньозваженого значення, яке має різні значення для оцінки глибини та покращення глибини,  $W$  – вагове ядро, яке визначається наступним рівнянням:

$$W(l, k) = e^{-\frac{(I(k)-I(l))^2}{2\delta_c^2}} \times e^{-\frac{(l-k)^2}{2\delta_d^2}}, \quad (1.15)$$

$N_k$  представляє локальне вікно з центром в  $k$ ,  $e'_k$  – відфільтроване значення,  $\delta_c$  та  $\delta_d$  контролюють смугу пропускання ядер дальності та відстані відповідно,  $I$  – кольорове зображення, а  $I(k)$ ,  $I(l)$  – його елементи,  $k$ ,  $l$  – двовимірне розташування пікселів на зображенні.

По суті, цей фільтр є анізотропним середньозваженим. Вага для кожної пари пікселів обчислюється за різницею кольорів та евклідовою відстанню, які відповідають двом частинам у рівнянні (1.15). Зокрема, це базується на розумному припущенні, що елементи, які відповідають пікселям, схожим за кольором та розташуванням з центральним пікселем, мають більшу вагу в середньому.

Для згладжування зі збереженням країв,  $e_l$  є значенням для пікселя  $l$  у довільному каналі [34]. Це може запобігти появі пікселів, які дуже відрізняються за кольором, а центральний піксель бере участь в усередненні. Таким чином, краї можуть бути збережені під час згладжування зображення.

Для агрегування витрат при стерео співставленні,  $e_l$  – це вартість збігу пікселів з певним зміщенням  $d$  у локальному вікні [35]. Остаточна вартість збігу є середньозваженим значенням цих вартостей збігу пікселів. Передбачається, що пікселі, які схожі за кольором з центральним пікселем, мають схожі значення зміщення, і вони повинні відігравати домінуючу роль у середньозваженому значенні. Ця схема також називається м'якою сегментацією [36], яку можна визначити за допомогою такого рівняння:

$$C(k, d) = \frac{\sum_{l \in N_k} W_{ref}(l, k) W_{tar}(l - d, k - d) diff(l, l - d)}{\sum_{l \in N_k} W_{ref}(l, k) W_{tar}(l - d, k - d)}, \quad (1.16)$$

де  $W_{ref}$  і  $W_{tar}$  – ваги, обчислені двостороннім фільтром (рівняння (1.15)) для еталонного і цільового вікон відповідно,  $diff$  – вартість співставлення на рівні пікселів. Знаменник – коефіцієнт нормалізації.

Для надроздільної реконструкції карти глибини, автори роботи [37] запропонували спільну двосторонню дискретизацію (JBU), краї карти глибини низької роздільної здатності можуть бути уточнені відповідно до країв зареєстрованого кольорового зображення високої роздільної здатності,  $e_l$  – нефільтроване значення глибини для пікселя  $q$ ,  $e'_k$  – фільтроване значення для пікселя  $k$ .

На основі цієї роботи було запропоновано кілька варіантів. Автори роботи [38] пропонують фільтр, який обчислює вагові коефіцієнти на основі геодезичних даних. Геодезичні дані обчислюються в спільному просторі кольору і відстані замість того, щоб розділяти простір кольору і простір відстані. Порівняно з JBU, спільне геодезичне ядро інтегрує зміни кольору вздовж геодезичних кривих, тому воно більш чутливе до тонких структур і дрібних змін масштабу, створюючи гладкі поверхні з різкими межами оклюзії.

Автори роботи [39] пропонують керовану фільтрацію зображень для покращення глибини, яка моделює лінійну залежність між вихідним і направляючим зображенням. Вона ґрунтується на припущенні, що вихідний сигнал має край тільки тоді, коли вхідний сигнал має край. Для визначення лінійних коефіцієнтів шукається рішення, яке мінімізує функцію енергії. Вона формується із загальної квадратичної різниці між відфільтрованим зображенням і вхідним зображенням з параметром регуляризації для запобігання занадто великих коефіцієнтів. Всі ці методи, включаючи JBU, мають ту саму форму з різними ваговими ядрами, що й рівняння (1.14). Ядра в роботі [38] та [39] визначаються як рівняння (1.17) і (1.18) відповідно:

$$W(l, k) = e^{-\frac{G_d(l, k)}{2\delta^2}}, \quad (1.17)$$

$$W(l, k) = \frac{1}{|\omega|^2} \sum_{p: (k, l) \in N_p} \left( 1 + \frac{(I(k) - \mu_p)(I(l) - \mu_p)}{\sigma_p^2 + \epsilon} \right), \quad (1.18)$$

де  $G_d(l, k)$  – геодезична відстань між  $l$  та  $k$ ,  $\mu_p$  – середнє значення локального вікна з центром у  $p$ ,  $p: (k, l) \in N_p$  позначає всі вікна, які включають  $k$  та  $l$ ,  $|\omega|$  – кількість пікселів у локальному вікні  $N_p$ ,  $\epsilon$  – мале значення для запобігання того, щоб знаменник не перетворився на нуль.

*Фільтри оптимізації за L1-нормою.* Ядра, представлені вище, базуються на оптимізації за L2-нормою. В якості альтернативи, мотивованої агрегуванням витрат при стерео співставленні, запропоновано деякі підходи, засновані на мінімізації за L1-нормою, яка є більш стійкими до викидів, ніж мінімізація за L2-нормою. Автори роботи [40] пропонують метод надроздільну реконструкцію карти глибини, заснований на методах спільної двосторонньої фільтрації (JBF) з набором кандидатів на глибину для ітеративного уточнення карти глибини високої роздільної здатності.

Остаточне значення глибини обирається за допомогою методу «переможець отримує все» (WTA) на основі обсягу витрат після попередньо визначеної кількості ітерацій. Експеримент показує, що цей метод може дати кращі результати зі збереженням країв. Автори роботи [41] пропонують метод фільтрації на базі зваженого підходу на основі спільної гістограми. В процесі генерації гістограми обчислюється вага, заснована на схожості кольору між еталонним пікселем та його сусідніми пікселями на кольоровому зображенні. Ця вага використовується для підрахунку кожного біна на спільній гістограмі. Остаточне рішення визначається шляхом пошуку глобальної моди на гістограмі. Вони мають спільну форму таких рівнянь:

$$d_k^* = \arg \max_{d \in [d_{min}, d_{max}]} C(k, d), \quad (1.19)$$

$$C(k, d) = \sum_{l \in N_k} W(l, k) Er(d - d_l), \quad (1.20)$$

де  $C$  – обсяг витрат, під певною міткою  $d$  значення вартості для кожного пікселя  $k$  обчислюється як середньозважене,  $W(l, k)$  має той самий зміст, що і рівняння (1.15),  $Er$  – функція похибки, яка є відмінністю між методами цього типу.

З цих рівнянь видно схожість між фільтром такого типу та агрегуванням витрат при стерео співставленні рівняння (1.5). Знову ж таки, вони поділяють те саме припущення, що пікселі, схожі за кольором, мають схоже значення глибини.

Автори робіт [41, 42] будують функцію похибки як наступні рівняння:

$$Er(d - d_l) = \min(|(d - d_l)|, \tau), \quad (1.21)$$

$$Er(d - d_l) = e^{-\frac{(d-d_l)^2}{2\delta^2}}, \quad (1.22)$$

де  $\tau$  і  $\delta$  є наперед визначеними параметрами. Всі локальні методи можна виконувати ітеративно. Загалом, складність локальних методів є низькою, тому їхня ефективність є привабливою. Однак вони завжди поступаються глобальним аналогам, особливо у випадку зашумлення.

### 1.3.5 Глобальна оптимізація та графічні моделі

*Зв'язок байєсівським виведенням.* Порівняно з локальними методами, глобальні аналоги є більш надійними. Конкретна задача моделюється як задача оптимізації заздалегідь визначеної цільової функції. Таку функцію також називають функцією енергії. Майже всі глобальні методи належать до байєсівського виведення. Наприклад, загальна модель для марковського випадкового поля (MRF) визначається як наступне рівняння згідно з теоремою Хаммерслі-Кліффорда [43]:

$$pb(X|Y) = \frac{pb(Y|X)pb(X)}{\sum_X pb(Y|X)pb(X)}, \quad (1.23)$$

$$pb(Y|X) = e^{-data(X,Y)}, \quad (1.24)$$

$$pb(X) = e^{-\lambda reg(X)}, \quad (1.25)$$

де  $X$  – результат виведення (тобто, високоякісна карта глибини для покращення або оцінки глибини),  $Y$  – спостережувана інформація для висновку (наприклад, спостережувані значення глибини для покращення глибини, помилка узгодження для оцінки глибини),  $pb$  – аббревіатура для ненормованої ймовірності,  $data(X, Y)$  – член даних у функції енергії, який описує ймовірність появи певного стану  $Y$  при заданому  $X$ ,  $reg(X)$  – член регуляризації, який моделює попередній розподіл самого  $X$ ,  $\lambda$  – коефіцієнт збалансованості між ними.

Виходячи з теорії максимальної апостеріорної оцінки, висновок може бути розв'язаний за формулою  $\arg \max_X pb(X|Y)$ . Оскільки знаменник пов'язаний лише з  $Y$ , його можна проігнорувати. Отже, еквівалентним рішенням є максимізація  $\widehat{pb}(X|Y)$ , визначеної як наступне рівняння:

$$X^* = \arg \max_X \widehat{pb}(X|Y) = \arg \max_X pb(Y|X)pb(X). \quad (1.26)$$

Виходячи з цього, мінімізація функції енергії  $E$ , яка визначається як наступне рівняння, еквівалентна максимізації  $\widehat{pb}(X|Y)$ :

$$X^* = \arg \max_X E(X, Y) = data(X, Y) + \lambda reg(X), \quad (1.27)$$

Член регуляризації  $reg$  відповідає певному пріоритету. Наприклад, відповідно до попередньої умови згладжування,  $reg$  можна змоделювати як наступне рівняння:

$$pb(X) = e^{-\lambda_{reg}(X)} = e^{-\|Vec(\nabla X)\|_2}, \quad (1.28)$$

де  $\nabla$  – градієнтний оператор, який обчислює різницю між кожним пікселем та його сусідніми пікселями,  $Vec$  – векторизований оператор, який перетворює всі градієнти у вектор.

*Байєсівське виведення для оцінки глибини та покращення карти глибини.* Для оцінки глибини,  $X$  та  $Y$  – це глибина  $D$  та спостережувану вартість  $C(D)$  для всіх пікселів відповідно. Зокрема, для певного пікселя  $k$  істинна глибина  $d_k \in D$  з більшою ймовірністю відповідає найнижчій вартості  $C(d_k)$ . Таким чином,  $pb(C(D)|D)$  можна змоделювати як наступне рівняння:

$$pb(C(D)|D) = e^{-data(D)} = e^{-C(D)}. \quad (1.29)$$

Це означає, що чим вища вартість, тим нижча ймовірність того, що буде обрано справжню глибину.

Для покращення карти глибини,  $X$  і  $Y$  – це виведена глибина  $D$  і спостережувані значення глибини  $O$  відповідно. Значення  $pb(O|D)$  пов'язане з різницею між  $D$  і спостережуваною глибиною  $O$ , що показано у вигляді наступного рівняння:

$$pb(O|D) = e^{-data(D,O)} = e^{-f(D,O)}. \quad (1.30)$$

Оціночна функція  $f$  може бути спеціально визначена (наприклад, це може бути середньоквадратична похибка). Вважається, що виведені значення глибини повинні узгоджуватися зі спостережуваними.

*Дискретна оптимізація за допомогою розрізу графа.* Оскільки оцінка глибини полягає в оцінці дискретних змінних зміщення, вона є задачею дискретної оптимізації. Після квантифікації реального значення глибини, покращення глибини також може бути задачею дискретної оптимізації. Розріз графа [44] є популярним

алгоритмом дискретної мінімізації енергії на основі графа. Щоб спростити пояснення, моделюючи оцінку глибини та покращення карти глибини як задачу оптимізації графів, спочатку вводиться базове поняття розрізу графа.

Перше поняття полягає в тому, що розріз графа розв'язує бінарну дискретну задачу оптимізації шляхом перемаркування всіх змінних. Існує дві стратегії перепризначення: « $\alpha, \beta$  – заміна» та « $\alpha$ -розширення». Значення  $\alpha, \beta$  позначають можливі двійкові значення змінних. Стратегія « $\alpha, \beta$ -заміни» оптимально перемаркує пікселі, змінюючи мітки між  $\alpha$  та  $\beta$ . Стратегія « $\alpha$  – розширення» просто розширює множину пікселів з міткою  $\alpha$  для мінімізації функції енергії. Оскільки в дослідженні надається значна перевага стратегії « $\alpha$ -розширення» [45], далі буде розглядатися лише ця стратегія. Більш детально відмінності між різними стратегіями описані в роботі [46]. Така операція виконується безперервно, поки енергія не збігається. Початкові значення міток можуть бути наперед визначеними або випадковими. У цій частині окремі малі літери позначають пікселі, наприклад,  $k, l$ , мітки для них визначаються як  $d_k, d_l$ .

Друга концепція полягає в тому, що перемаркування виконується шляхом відсікання ребер, що з'єднують вузли, і двох кінцевих вузлів  $\alpha, \alpha'$ . Більш конкретно, коли перерізається ребро, що з'єднує вершину і  $\alpha$ , мітка цієї вершини змінюється на  $\alpha$ . Мітка вершини залишається незмінною, коли перерізається ребро, що з'єднує вершину і  $\alpha'$ .

Третя концепція полягає у тому, що правильний план розрізання полягає у розділенні двох кінцевих вершин  $\alpha, \alpha'$ . Це означає, що після розрізання між ними немає шляху. Для кожного плану розрізання, вартість є сумою ваг усіх ребер, що розрізаються. Серед цих рішень перевага надається тому, яке має найменшу вартість.

Для задачі оцінки глибини, мітки є кандидатами на зміщення в діапазоні пошуку. А мітки для задачі покращення карти глибини – це рівні квантування. Тому такі задачі є задачами оптимізації з декількома мітками і не можуть бути оптимізовані безпосередньо за допомогою розрізу графа. Альтернативою є

декомпозиція задачі з декількома мітками на багато підзадач двійкової дискретної оптимізації. Тоді вихідна задача наближено розв'язується шляхом розв'язування підзадач по порядку.

Далі треба проілюструвати моделювання розв'язання енергетичної функції рівняння (1.27) за допомогою розрізу графа, для вирішення задачі оцінки глибини. Граф складається з набору вершин, кожна з яких представляє піксель. Дві кінцеві вершини  $\alpha$ ,  $\alpha'$  з'єднані з усіма вершинами ребрами. Також існують ребра, що з'єднують вершини між собою. Вага ребер визначається членом даних та членом регуляризації. Таким чином, при певній конфігурації графа пошук оптимального плану розрізу може бути еквівалентним пошуку мінімуму функції енергії (1.27) згідно з третьою концепцією. Для спрощення пояснень, необхідно зосередитися на підзадачі оптимізації бінарного маркування, побудованої на графі з двома вершинами  $k, l$ .

На рисунку 1.9 наведено ілюстрацію розрізу графа.

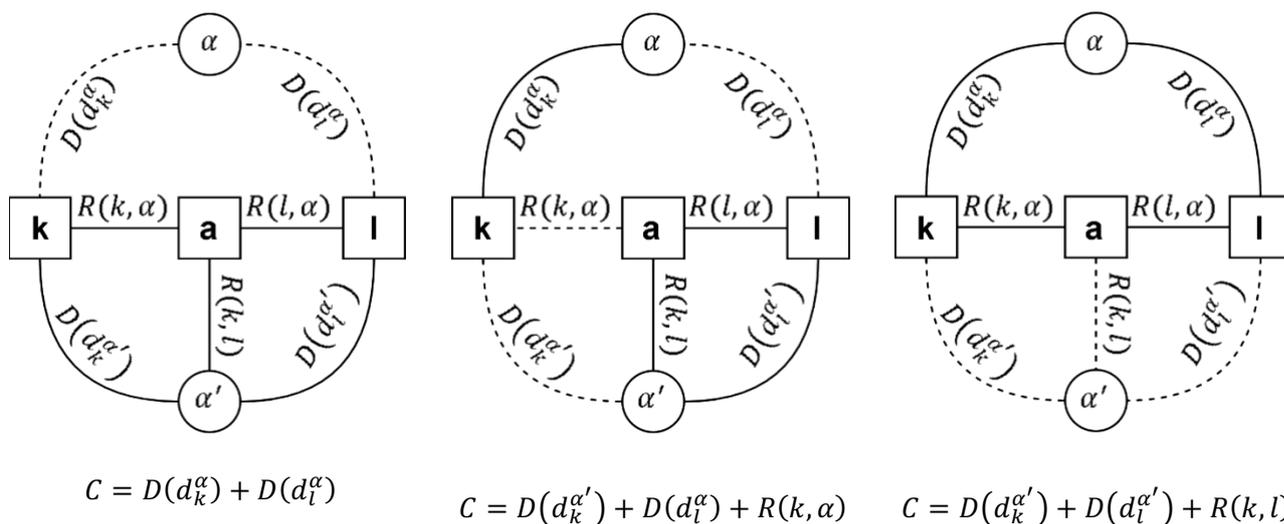


Рисунок 1.9 – Ілюстрація розрізу графа з трьома стратегіями розрізу

На рисунку ребра, показані пунктирними лініями, обрізані, а для кожного рішення проілюстровано вартість.  $D$  та  $R$  є ваговими функціями для ребер, що представляють член даних та член регуляризації. Ребра, що з'єднують вершини з термінальними вершинами, називаються ребрами члену даних, наприклад,  $d_k^\alpha$

позначає ребро, що з'єднує вершину  $k$  з термінальною вершиною  $\alpha$ , причому  $k$  мав мітку  $d_k$  перед перемаркуванням. А ребра, що з'єднують вершини з іншими вершинами, називаються ребрами члену регуляризації, наприклад,  $(k, l)$  позначає ребро, що з'єднує вершину  $k$  з вершиною  $l$ .

Для вершин з міткою  $d_k$ , відмінною від  $\alpha$  до переміщення, ребра, що з'єднують їх з  $\alpha$ , зважуються з вартістю, отриманою з функції вартості. Для вершин, що вже мають мітку  $\alpha$ , ваги ребер, що з'єднують їх з  $\alpha$ , дорівнюють нескінченності, щоб гарантувати, що ці вершини зберігають свою мітку  $\alpha$  протягом всієї оптимізації. На основі поведінки ребер даних, що обрізаються, описаної вище, вводяться допоміжні вершини (наприклад, вершина  $a$  на рисунку 1.9) для моделювання умови регуляризації в ситуаціях, коли ребра даних, що обрізаються, з'єднують вершину та її сусідні вершини з усіма варіантами термінальних вершин  $\alpha, \alpha'$  (наприклад,  $k, l$  на рисунку 1.9). Відображення ваг ребер у графі у функцію енергії (1.27) наведено в таблиці 1.2.

Таблиця 1.2 – Ваги ребер для прикладу на рисунку 1.9

Ребро	Вага	Стан перед переміщенням
$D(d_k^{\alpha'})$	$inf$	$d_k = \alpha$
$D(d_k^{\alpha'})$	$data(d_p)$	$d_k \neq \alpha$
$D(d_k^{\alpha})$	$data(\alpha)$	Нема
$R(k, a)$	$\lambda reg(d_k, \alpha)$	$l \in N_k$
$R(a, l)$	$\lambda reg(\alpha, d_k)$	$l \in N_k$
$R(k, l)$	$\lambda reg(d_k, d_l)$	$l \in N_k, d_k \neq d_l$
$R(k, l)$	0	$l \in N_k, d_k = d_l$

При заданій конфігурації графа мінімізація функції енергії дорівнює знаходженню мінімального розрізу графа, що здійснюється шляхом обчислення максимального потоку між термінальними вершинами. Алгоритм максимального потоку представлений в роботі [47]. Алгоритм розрізу графа здатен точно

розв'язувати задачі оптимізації бінарного маркування. Для задач оптимізації мультимаркування, таких як оцінка глибини та покращення карти глибини, він дає лише наближений розв'язок, який є локальним мінімумом, близьким до глобального мінімуму.

### 1.3.6 Надроздільна реконструкція карти глибини за допомогою методів на базі навчання

*Розріджене кодування. Модель синтезу.* Одне з припущень, яке виявилось успішним при реконструкції зображень, полягає в тому, що природні зображення допускають розріджене представлення  $\alpha \in R^m$  над певним надповним словником  $U \in R^{n \times m}$  з  $m > n$ . Вектор  $\alpha$  є розрідженим, коли більшість його коефіцієнтів дорівнюють або близькі до нуля. Коли ділянка сигналу або векторизованого зображення  $x$  має розріджене представлення  $\alpha$  над  $U$ , вона може бути виражена як лінійна комбінація лише декількох стовпців словника  $\{U_i\}_{i=1}^m$ , які називаються атомами, що показано у наступному рівнянні:

$$x = U\alpha. \quad (1.31)$$

Словник зазвичай вивчається за допомогою наступного рівняння:

$$U^*, \alpha^* = \arg \min_{U, \alpha} \|x - U\alpha\|_2^2 \text{ за умови } g(\alpha) \leq \epsilon, \quad (1.32)$$

де  $g$  – функція оцінки розрідженості,  $\epsilon \in R^+$  – оцінка верхньої межі розрідженості. В ідеалі, функція оцінки розрідженості повинна бути 0-нормою, тобто  $g(\alpha) = \|\alpha\|_0$ . Однак, оскільки її важко обчислити, замість 0-норми використовують  $p$ -норму ( $0 < p < 1$ ). На додаток до  $p$ -норми існує багато інших функцій розрідженості, наприклад, функція логарифма квадрата, яка використовується в роботі [48].

На основі вивченого словника та спостережуваного низькоякісного сигналу  $y$ , деградованого за рівнянням (2.13), оцінку вихідного сигналу можна отримати, спочатку розв'язавши розріджені коефіцієнти:

$$\alpha^* = \arg \min_{\alpha} g(\alpha) \text{ за умови } \|HU\alpha - y\|_2^2 \leq \epsilon, \quad (1.33)$$

а потім синтезувавши сигнал за обчисленими розрідженими коефіцієнтами за допомогою  $x^* = U\alpha^*$ . Існують репрезентативні роботи для надроздільної реконструкції карти глибини на основі моделі синтезу. В роботі [49] спільно навчають словники для зареєстрованих ділянок карт глибини низької роздільної здатності, карт глибини високої роздільної здатності та кольорових зображень.

На етапі реконструкції, карти глибини високої роздільної здатності реконструюються за допомогою розрідженого представлення вивченого відповідного словника. Автори роботи [50] запропонували метод, який використовує навчальний набір високоякісних кольорових зображень і переносить його інформацію на карту глибини низької роздільної здатності за допомогою багатомасштабного навчання словника.

Він вивчає словник геометричних примітивів, який фіксує кореляцію між високоякісними даними сітки, картами глибини низької роздільної здатності і кольоровими зображеннями високої роздільної здатності. Ці два методи мають спільну форму на етапі навчання:

$$U^*, \alpha^* = \arg \min_{U, \alpha} \left\| \begin{bmatrix} x_h \\ \tilde{x}_l \\ x_c \end{bmatrix} - \begin{bmatrix} U_h \\ U_l \\ U_c \end{bmatrix} \alpha \right\|^2 + \lambda \|\alpha\|_1, \quad (1.34)$$

де  $x_h$ ,  $\tilde{x}_l$  і  $x_c$  – характеристики для ділянок глибини високої роздільної здатності, ділянок глибини низької роздільної здатності і кольорових ділянок високої роздільної здатності відповідно;  $U_h$ ,  $U_l$  та  $U_c$  – словники для глибинних ділянок високої роздільної здатності, глибинних ділянок низької роздільної

здатності та кольорових ділянок високої роздільної здатності відповідно;  $\alpha$  – коефіцієнти розрідженості. Це ітеративна схема підвищення вибірки, яка вводить вимірювання нормалізованого коефіцієнта кореляції для прогнозування того, які кольорові границі з найбільшою ймовірністю збігаються з границями глибини.

*Модель аналізу.* На відміну від моделі синтезу (1.31), фаза реконструкції для моделі аналізу полягає у розв'язанні наступного рівняння:

$$x^* = \arg \min_x g(\Omega x) \text{ за умови } \|Hx - y\|_2^2 \leq \epsilon, \quad (1.35)$$

Тут  $\Omega \in R^{k \times n}$  при  $k > n$  називається оператором аналізу, а аналізований вектор  $\Omega x \in R^k$  вважається розрідженим, де розрідженість знову ж таки вимірюється за допомогою відповідної функції  $g$ . На відміну від моделі синтезу, де сигнал повністю описується ненульовими елементами  $\alpha$ , нульові елементи аналізованого вектора  $\Omega x$  описують підпростір, що містить сигнал в моделі аналізу. Щоб підкреслити цю різницю, введено термін со-розрідженості, який просто підраховує кількість нульових елементів  $\Omega x$  [51]. Різниця між цими двома типами моделей також проявляється на етапі навчання. Навчання оператора в моделі аналізу визначається за допомогою наступного рівняння:

$$x^*, \Omega^* = \arg \min_{x, \Omega} g(\Omega x) + \text{reg}(\Omega) \text{ за умови } \|x - y\|_2^2 \leq \epsilon, \quad (1.36)$$

де  $\text{reg}(\Omega)$  є додатковим обмеженням на оператор аналізу для уникнення тривіальних рішень. Дійсно, якщо на  $\Omega$  не накладено такої регуляризації, то помічено, що тривіальний розв'язок  $\Omega \equiv 0$  є глобальним мінімізатором. Такі регуляризації на операторі аналізу досліджено в роботі [52]. Далі наведені деякі з широко використовуваних обмежень:

– обмеження на норму рядка; всі рядки  $\Omega$  мають однакову норму, тобто  $\|\Omega_i\|$  для  $i$ -го рядка оператора  $\Omega$ ;

– обмеження на повний ранг; оператор аналізу  $\Omega$  має повний ранг, тобто  $rank(\Omega) = n$ ;

– обмеження на жорсткі рамки; допустима множина цього обмеження є множиною жорстких рамок в  $R^{k \times n}$ , тобто  $\Omega^T \Omega = I_k$ , де  $I_k$  є оператором тотожності в  $R^k$ .

На основі моделі аналізу, автори роботи [48] використовують розрідженість операторів аналізу, які застосовуються до карти глибини та відповідного кольорового зображення, і реконструюють карту глибини високої роздільної здатності через обмеження точності даних та розрідженості, керованої кольором. Оптимізовані задачі на етапах навчання та реконструкції показані у вигляді наступних рівнянь:

$$(\Omega_I^*, \Omega_D^*) \in \arg \min_{\Omega_I, \Omega_D \in OB(n,p)} g(\Omega_I x_I, \Omega_D x_D) + reg(\Omega_I) + reg(\Omega_D), \quad (1.37)$$

$$x_D^* \in \arg \min_{x_D} \lambda g(c, \Omega_D x_D) + \|H_D x_D - y_D\|_2^2, \quad (1.38)$$

де  $\Omega_I^*$  та  $\Omega_D^*$  – оператори аналізу для ділянок інтенсивності високої роздільної здатності і ділянок глибини високої роздільної здатності відповідно,  $x_I$  та  $x_D$  – ділянка зареєстрованого зображення інтенсивності і ділянка карти глибини наземної істини,  $g$  – функція розрідженості, оператор аналізу обмежений до  $OB(n,k)$ , який є множиною матриць повного рангу з нормалізованими стовпчиками,  $reg$  вказує на попередні знання про такі оператори аналізу,  $c$  – константа, яка означає, що аналізований сигнал інтенсивності не пов'язаний з  $x_D$ ,  $x_D$  – це відрізок глибини високої роздільної здатності,  $H_D$  – матриця деградації,  $y_D$  – множина спостережуваних значень глибини.

### 1.3.7 Використання вейвлет-перетворення при побудові карти глибин

Як було показано вище, існує низка рішень, які дають змогу створювати карти глибин за допомогою локального зіставлення пікселів та/або блоків [37-39].

У базовому методі обчислення карти диспаратності (величини відмінності в локалізації відповідних елементів зображень, отриманих із правої та лівої камери) виконується шляхом зіставлення пікселів на лівому та правому зображеннях для обчислення карти диспаратності (реалізовано у функції StereoBM з бібліотеки Open CV).

У роботі [53] показано, що для підвищення швидкодії та зниження енергоспоживання, запроповано обчислювати диспаратність для фрагмента рядка (блоку) на лівому зображенні, визначаючи, на скільки пікселів зміщене відповідне праве зображення.

Ідея пропонованого підходу полягає в наступному:

- для фрагменту рядка яскравості визначається область максимуму яскравості (виділення контуру об'єкту) на зображенні однієї з камер, наприклад, лівої (1) (рис. 1.10);

- сформується шаблон огинаючої яскравості: значення яскравості, відповідне координаті екстремуму рядка, і значення яскравості сусідніх пікселів та знаходиться схожий фрагмент на зображенні правої камери (2);

- знаходиться відстань між схожими елементами і відповідне йому значення диспаратету;

- обробку проводять за всіма рядками.

Такий підхід виявляється простішим у реалізації та працює швидше ніж методи, описані у розділі 1.3.2.

Обчислювальну складність знаходження шаблонів у запропонованому методі можна знизити завдяки тому, що для знаходження шаблонів зможна використовувати методику оцінювання глибини шляхом виділення контурів об'єктів із регульованою деталізацією за допомогою гіперболічного вейвлет-перетворення (ВП) [54].

Важливим етапом запропонованого підходу є також локалізація області екстремуму за допомогою вейвлет функції Гаара [55].

Для виділення контурів об'єктів (визначення області максимуму яскравості) за допомогою вейвлет-перетворення видаляється підсумовуюча складова

апроксимації (по вертикалі) найменшого масштабу ( $a_1$ ) дискретного вейвлет-перетворення шляхом прирівнювання її значень до нуля. Обчислення оберненого перетворення з використанням деталізуючої складової ( $d_1$ ) призводить до виділення контурів (по горизонталі) на реконструйованому таким чином зображенні. Процедуру виділення контурів за допомогою вейвлет-перетворення можна представити наступним чином:

$$\begin{aligned} f_0 &\rightarrow a_1 \rightarrow 0, \\ d_1 &\rightarrow f \rightarrow K(f'), \end{aligned} \quad (1.39)$$

де  $f_0$  – рядок (стовпець) вихідного зображення;  
 $a_1$  – підсумовуюча складова вейвлет-перетворення;  
 $d_1$  – деталізуюча складова вейвлет-перетворення;  
 $f'$  – рядок (стовпець) зображення після оберненого перетворення;  
 $K$  – оператор для отримання контуру.

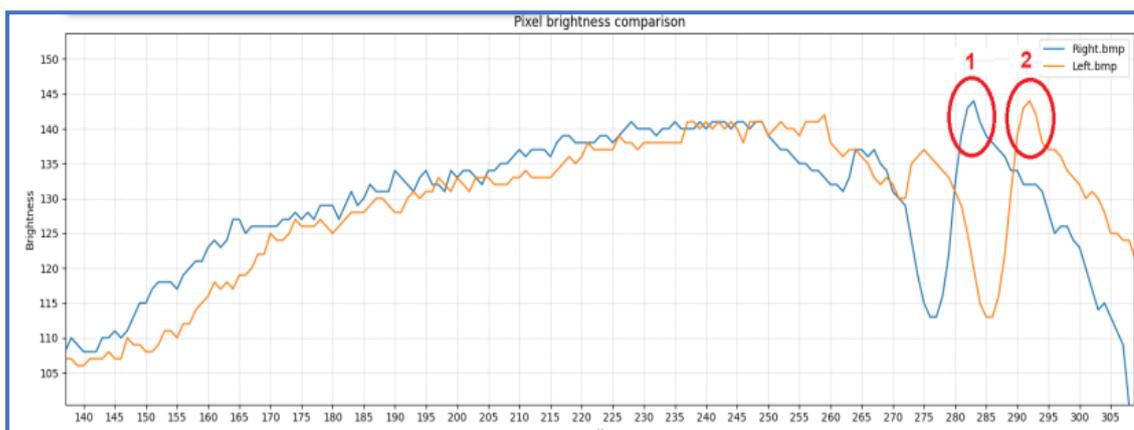


Рисунок 1.10 – Фрагменти рядка яскравостей правого та лівого кадру

#### 1.4 Шляхи удосконалення побудови карти глибин для систем технічного зору

Забезпечення високої точності та надійності побудови карти глибин є критичним аспектом для розвитку інтелектуальних систем технічного зору.

Незважаючи на наявність великої кількості існуючих методів, багато з них не забезпечують достатньої якості в умовах обмеженого освітлення, динамічних змін сцени або при наявності слабо текстурованих поверхонь.

У зв'язку з цим виникає потреба у розробці нових підходів, які здатні подолати наявні обмеження та розширити функціональні можливості систем технічного зору.

Швидке отримання карти глибини.

Порівняно з глобальними методами отримання карти глибини, локальні методи [56] мають низьку складність. Тому для додатків, що працюють у реальному часі, є сенс зосереджуватися на покращенні продуктивності локальних методів. Серед таких методів найкращих результатів завжди досягають методи на основі адаптивної ваги.

Вага показує, наскільки ймовірно, що поточний піксель має подібну розбіжність з центральним пікселем локального вікна.

Більшість методів на основі адаптивної ваги припускають, що така вага залежить від різниці кольорів і відстані між поточним пікселем і центральним пікселем.

На основі таких ваг можна адаптивно керувати внеском кожного сусіднього пікселя у відповідну агреговану вартість співставлення.

Незважаючи на гарні результати, які дають методи на основі адаптивної ваги, обчислення ваги для кожного пікселя та оцінка глибини є дуже трудомісткими завданнями.

Ці процедури є обчислювально надлишковими, оскільки карти глибини завжди є гладкими. В рамках дисертаційного дослідження пропонується використання атрибуту гладкості замість попівсельного оцінювання глибини, що дозволило б значно зменшити час роботи методу.

*Покращення часової узгодженості для послідовних карт глибини.*

Послідовності карти глибини, згенеровані в процесі оцінювання глибини для кадрів відеопотоку, страждають від проблеми часової неузгодженості. В ідеалі, значення глибини статичних об'єктів повинні залишатися однаковими в сусідніх

кадрах, але часто вони оцінюються як різні значення. Ці часові помилки глибини значно погіршують візуальну якість синтезованого віртуального представлення, а також ефективність кодування послідовних карт глибини [57]. В даному дисертаційному дослідженні пропонується метод покращення часової узгодженості.

#### *Вимірювання неузгодженості країв між картою глибини та кольоровим зображенням*

Серед керованих методів покращення глибини без використання зовнішніх наборів даних, існує низка методів [58-60], які намагаються зменшити артефакти копіювання текстури та розмиття країв глибини шляхом балансування внеску від оригінальної карти глибини та супутнього кольорового зображення. Однак, ці методи не оцінюють в явному вигляді неузгодженість країв між кольоровим зображенням і відповідною картою глибини.

Тому вони не можуть адаптивно контролювати величину впливу кольорового зображення на процес покращення карти глибини, що призводить до появи артефактів копіювання текстури та розмиття країв глибини. В рамках даного дисертаційного дослідження представлені два методи вимірювання неузгодженості країв на базі жорсткого та м'якого прийняття рішень. Запропоновані методи дозволяють отримати більш точне визначення невідповідності в числовому вигляді.

#### *Вдосконалене покращення карти глибини за допомогою попередньо визначеної моделі*

Ідея вдосконалення керованого методу покращення карти глибини полягає у його вбудовуванні у попередньо визначену модель, наприклад, Марковське випадкове поле (МВП). Це дозволяє у значній мірі уникнути появи артефактів копіювання текстури. Крім того, в існуючих методах, які використовують Марковське випадкове поле, член регуляризації обчислюється лише на основі різниці в кольорі та глибині між пікселем та його сусідніми пікселями на кольоровому зображенні та грубо інтерпольованій карті глибини, відповідно. Така схема обчислень ігнорує локальну структуру карти глибини і в даному

дисертаційному дослідженні називається неструктурною схемою. Тому, у випадку використання великого коефіцієнта дискретизації, такі методи генерують надмірно згладжені краї глибин. Для збереження структури на картах глибин дуже низької якості потрібна більш стійкий метод, який представлений в даному дисертаційному дослідженні.

## 1.5 Висновки до першого розділу

У першому розділі дисертації було проведено ґрунтовний аналіз актуальності та сучасного стану проблеми побудови карт глибин для систем технічного зору. Досліджено поняття систем технічного зору, їх основні характеристики, класифікацію, галузі застосування та тенденції розвитку. Підкреслено, що використання карт глибин суттєво розширює можливості комп'ютерного зору, забезпечуючи тривимірне розуміння сцени, яке є критично важливим для задач робототехніки, безпеки, автономного транспорту та інших галузей.

В ході огляду були виявлені основні технології отримання карт глибин, серед яких пасивні (стереозір) та активні методи (датчики ToF та структуроване світло). Встановлено, що існуючі методи мають певні недоліки, такі як низька роздільна здатність, шумність, наявність прогалин та артефактів. Крім того, аналіз літератури виявив суттєві обмеження при використанні традиційних алгоритмів стереоспівставлення, особливо у випадках оклюзій та безтекстурних ділянок.

Окрему увагу приділено підходам удосконалення якості карт глибин, зокрема методам, що використовують інформацію з супутніх кольорових зображень (керовані методи). Виявлено, що керовані методи, зокрема із застосуванням алгоритмів на основі штучного інтелекту, демонструють кращі результати, але також мають свої проблеми, такі як артефакти копіювання текстури та розмиття країв.

Таким чином, отримані результати аналізу підтверджують необхідність подальших досліджень і розробки нових методів побудови та покращення карт глибин, які були б більш ефективними, надійними і менш чутливими до складних

умов експлуатації. Запропоновані напрями удосконалення повинні зосереджуватися на інтеграції різних технологій, використанні новітніх алгоритмів машинного навчання та оптимізації існуючих методик з урахуванням практичних вимог і специфіки застосування систем технічного зору.

Сформульовано мету та задачі дослідження:

**Метою дослідження** є підвищення ефективності побудови карти глибин для систем технічного зору шляхом удосконалення пасивних та активних методів побудови карт глибин.

Для досягнення мети дослідження необхідно вирішити наступні завдання:

– провести аналіз основних проблем, обмежень та викликів, які виникають в процесі побудови карт глибин в системах технічного зору;

– удосконалити пасивний локальний метод побудови карти глибини за рахунок використання процедури адаптивного зіставлення та афінно-інваріантних ознак;

– удосконалити метод покращення часової узгодженості для послідовностей карт глибини за рахунок впровадження показника надійності глибини та використання динамічних атрибутів регіонів;

– удосконалити керований метод покращення точності карти глибини за рахунок впровадження процедури оцінки неузгодженості країв вихідного кольорового зображення та отриманої карти глибини з жорстким та м'яким прийняттям рішень;

– удосконалити керований метод покращення точності карти глибини низької якості на базі Марківського випадкового поля за рахунок впровадження процедури оцінки неузгодженості країв з м'яким прийняттям рішень та використанням мінімального кістякового лісу;

– розробити інструментальні засоби, що реалізують запропоновані рішення та провести апробацію і впровадження отриманих теоретичних результатів.

## 2 ПІДВИЩЕННЯ ЕФЕКТИВНОСТІ ПОБУДОВИ КАРТИ ГЛИБИНИ ЗА ДОПОМОГОЮ ПАСИВНИХ МЕТОДІВ

### 2.1 Метод побудови карти глибини з використанням адаптивної процедури зіставлення та афінно-інваріантних ознак

Карти глибин можна отримати за допомогою процедури стереозіставлення, використовуючи кольорові зображення, отримані з декількох ракурсів. Існуючі методи можна розділити на дві категорії: локальні методи та глобальні методи. Локальні методи мають низьку складність, вони обчислюють значення глибини для кожного пікселя незалежно шляхом зіставлення лише локальної інформації.

Порівняно з локальними методами, глобальні методи є складнішими, оскільки вони визначають значення глибини для всіх пікселів одночасно на основі попередньо визначених параметрів (наприклад, норми загальної варіації).

Як було зазначено в першому розділі даного дисертаційного дослідження, одним з базових етапів оцінки глибини є етап агрегування вартості. Агрегування вартості полягає у створенні множини, кожен елемент якого є відповідною вартістю між пікселем-еталоном і пікселем-ціллю, визначеним певним кандидатом на нерівність.

Для більшої надійності, агрегування витрат виконується в межах локального вікна для кожного пікселя. Для визначення таких локальних вікон було запропоновано кілька методів:

- методи на основі декількох вікон, які вибирають найкраще локальне вікно з набору заздалегідь визначених розмірів [61];

- методи на основі змінних вікон не вибирають найкраще вікно з наперед визначеного набору, а замість цього обчислюють оптимальне локальне вікно для кожного пікселя [62];

- методи на основі адаптивної ваги адаптивно контролюють вагу для кожного пікселя всередині великого вікна фіксованого розміру [63-64].

На відміну від чіткого визначення форми кожного пікселя в методах зі змінним вікном, методи на основі адаптивної ваги приймають рішення неявно, що вводить поняття нечіткості.

Як було показано в декількох роботах, серед цих трьох типів локальних методів методи на основі адаптивної ваги завжди досягають найкращих результатів [35].

Значення ваги відповідає ймовірності того, що піксель має подібну розбіжність з центральним пікселем локального вікна.

Більшість методів на основі адаптивної ваги припускають, що така вага визначається різницею кольорів і відстанню між поточним пікселем і центральним пікселем [36].

Таким чином можна адаптивно контролювати вклад кожного пікселя в локальному вікні в агрегації витрат.

Незважаючи на відносно задовільні результати, які дають методи на основі адаптивних ваг, обчислення ваг для кожного пікселя та оцінка глибини є дуже трудомісткими процедурами. Вони мають високу обчислювальну надлишковість, оскільки карти глибини завжди є гладкими.

В рамках даного дисертаційного дослідження запропоновано адаптивну схему зіставлення, яка дозволяє зменшити складність обчислень і досягти продуктивності, близької до сучасних локальних методів.

Крім того, хоча більшість методів стереозіставлення використовують інформацію про колір для зіставлення згідно з припущенням, що пікселі, які зіставляються, повинні мати однаковий колір, це неефективно в областях без текстури.

Відсутність візуальних ознак робить стереозіставлення складним завданням. Для вирішення проблеми зіставлення в таких регіонах, запропонований метод використовує афінно-інваріантні ознаки для отримання більш надійних результатів.

### 2.1.1 Афінно-інваріантні ознаки

Перед тим, як надати опис запропонованого методу, необхідно дати визначення афінно-інваріантним ознакам. Нехай в евклідовому просторі  $R^3$  задано пряму  $ab$ . Нехай  $d$  – точка, що лежить всередині прямої  $ab$ . Пряма проектується на дві площини  $K$  і  $M$  з відповідними проекціями  $a_K, b_K, d_K$  і  $a_M, b_M, d_M$ , як показано на рисунку 2.1.

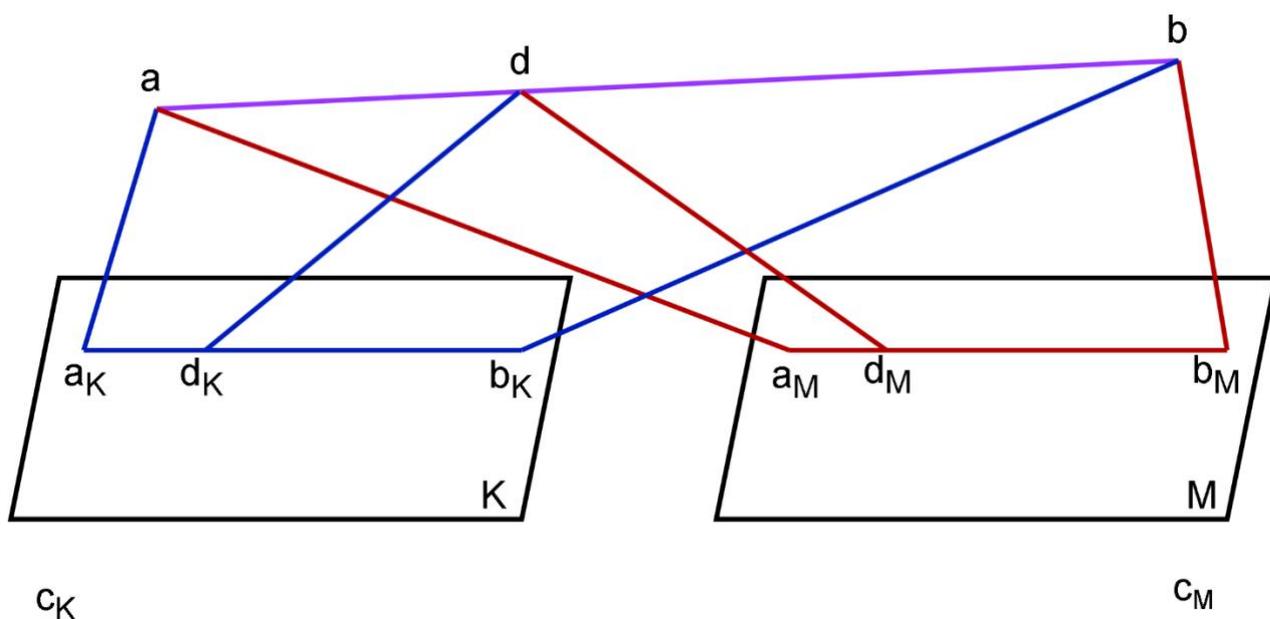


Рисунок 2.1 – Ілюстрація афінно-інваріантних ознак

Нехай  $c_K, c_M$  – оптичні центри площин проєкцій. Афінно-інваріантні ознаки можна змодельовувати за допомогою наступного рівняння [65]:

$$\frac{d_K - a_K}{b_K - a_K} \equiv \frac{d_M - a_M}{b_M - a_M}. \quad (2.1)$$

Щоб використовувати афінно-інваріантні ознаки для стерео зіставлення, спочатку слід обговорити дві проблеми. Перша полягає в тому, що афінно-інваріантні ознаки теоретично придатні для проєкції на плоску поверхню [16]. Однак реальна сцена складніша. Інша проблема полягає в тому, як знайти два

відповідні відрізки (наприклад,  $a_K b_K$  та  $a_M b_M$  на рисунку 2.1). Запропонований метод вирішує ці дві проблеми наступним чином:

1) припускається, що карта глибин складається з плоских поверхонь, як і в результаті роботи більшості сучасних методів на базі стереозіставлення [66]. Для Ламбертових поверхонь більша кривина також означає більшу варіацію через затінення, і ступінь цієї варіації інтенсивності може бути використаний для розбиття всієї кривої поверхні на ряд невеликих ділянок з низькою кривизною. Кожна ділянка може бути апроксимована до плоскої поверхні. Таким чином, відношення відстані у якості афінно-інваріантної ознаки може бути використано для неплоских, безтекстурних поверхонь;

2) добре відомо, що два відповідні відрізки знаходяться в одному рядку ліній розгортки, коли пари зображень випрямлені. У даному випадку вони виділяються шляхом пошуку вздовж ліній розгортки в лівому та правому напрямках з попередньо визначеним порогом різниці кольорів. Виходячи з наведеного вище аналізу, відношення відстані між лініями можна використовувати як афінно-інваріантну ознаку для процедури зіставлення.

Афінно-інваріантна ознака для кожного пікселя може бути визначена наступним чином:

$$F(p) = \frac{\sum_{n \in K_p} v_n}{\sum_{n \in K_p + M_p} v_n}, \quad (2.2)$$

де  $p$  – поточний піксель,  $K_p$  та  $M_p$  – множини пікселів, що складаються з пікселів від лівого крайнього пікселя та правого крайнього пікселя відрізка до  $p$  відповідно,  $n$  – сусідній піксель пікселя  $p$  на цьому відрізку. Для кожного пікселя  $p$  задається глобальний поріг для визначення  $K_p$  та  $M_p$ .

Пошук вздовж лівого та правого напрямків зупиняється до тих пір, поки абсолютна різниця яскравості двох сусідніх пікселів не стане більшою за заданий

пори́г. В ідеалі,  $v_n$  – це бінарне значення, яке вказує, чи належить даний піксель до цього відрізка, чи ні.

На додаток до жорсткого порогу пошуку для знаходження відрізка використовується схема м'якого прийняття рішення для перевірки ймовірності того, що кожен піксель  $n$  належить до цього відрізка. Ненормована ймовірність цього моделюється за наступним:

$$v_n = e^{-\frac{(L(p)-L(n))^2}{2\delta^2}}, \quad (2.3)$$

де  $L(p)$  та  $L(n)$  – яскравість пікселів  $p$  та  $n$  відповідно. Це означає, що чим менша різниця в яскравості, тим більша ймовірність того, що пікселі належать до одного відрізка, і навпаки. Після обчислення афінно-інваріантних ознак для всіх пікселів можна отримати спеціальне зображення, яке має таку ж роздільну здатність, як і кольорове зображення.

### 2.1.2 Початкове оцінювання глибини

Однією із проблем існуючих пасивних локальних методів отримання карти глибини є реалізація задачі агрегації вартості в межах локальних вікон, яка виконується на різних областях зображення. Якщо в області зображення відсутня текстура або вона не достатньо присутня, використання маленького фіксованого вікна не дозволить забезпечити достатню кількість ознак для подальшого зіставлення. З іншого боку, використання великого вікна в регіонах з текстурою призводить до отримання неточних меж глибини. Для подолання цієї проблеми, в запропонованому методі отримання карти глибини запропонована процедура адаптивного зіставлення, яка дозволяє класифікувати ситуації, описані вище.

Для безтекстурних областей адаптивні вікна генеруються на основі надмірної сегментації. Запропонований метод отримує еталонне значення глибини для всього регіону, після чого відбувається уточнення для кожного пікселя в межах

невеликого діапазону кандидатів на відмінність. Така обробка призводить до того, що в адаптивному локальному вікні з'являються пікселі, які мають схожий набір кольорів, що допомагає вирішити проблему зіставлення в безтекстурних регіонах [67]. Для областей з високим вмістом текстури, щоб отримати точні краї та зменшити час роботи, використовується невелике локальне вікно. Завдяки адаптивній схемі зіставлення, запропонований метод оцінки глибини може зменшити час роботи порівняно з сучасними локальними методами оцінки глибини. Структура запропонованого методу наведена на рисунку 2.2.

*Агрегування початкових витрат.* Першим кроком для оцінки глибини є агрегування витрат, яке виконується за допомогою запропонованої адаптивної схеми зіставлення.

Робота запропонованого методу полягає в отриманні зареєстрованої карти розбіжностей для кольорового зображення під певним кутом зору, яке називається еталонним зображенням, та зображеннями під іншими кутами зору називаються цільовими зображеннями.

Еталонне зображення сегментується на регіони з перекриттям за допомогою алгоритму сегментації зі зсувом середнього значення. Отримані регіони формують множину  $R(r_1, \dots, r_{max})$  [68].

На базі кількості пікселів в кожному регіоні  $Num(r_i)$ ,  $r_i$  може бути класифікований як «гладку область», якщо  $Num(r_i)$  більше задалегідь визначеного порогу, або на «негладку область» у іншому випадку. Для гладких областей, завдяки сегментації з перекриттям, більш імовірно задовольнити припущення, що розбіжності пікселів у кожній області є подібними.

Таким чином, для отримання вартості зіставлення при певній нерівномірності значень між усіма пікселями в цій області, використовується метрика суми абсолютних різниць, яка обчислюється за формулою:

$$SUM(m, cnd) = \sum_{p \in r_k} |ILM_{tg}(p + cnd) - ILM_{rf}(p)|, \quad (2.4)$$

де  $SUM(m, cnd)$  – помилка зіставлення  $m$ -го регіону для кандидата на нерівномірність  $cnd$ ;  $ILM$  – яскравість певного зображення;  $p + cnd$  – цільові пікселі в цільовому зображенні, що відповідають пікселю  $p$  на еталонному зображенні.

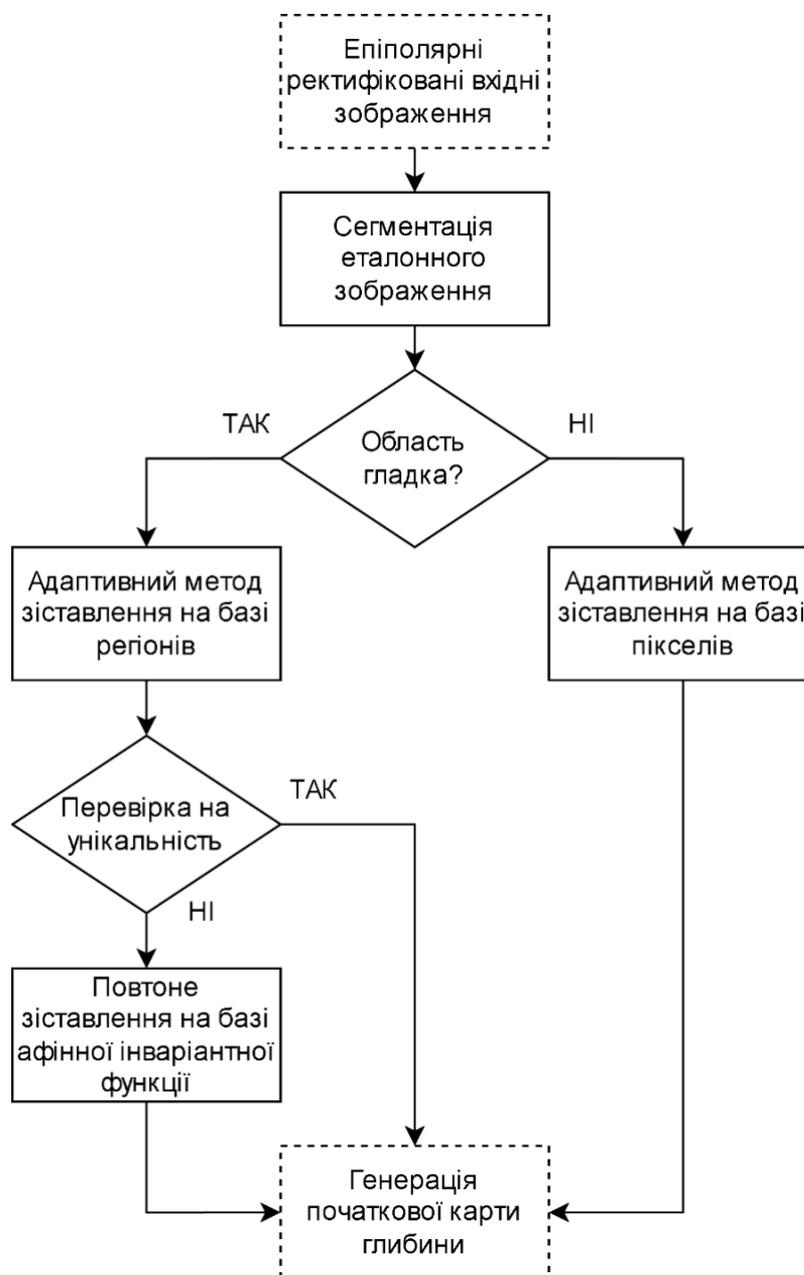


Рисунок 2.2 – Структура запропонованого методу

Для «негладких областей», агрегація вартості виконується для кожного пікселя в межах локального вікна. З огляду на обчислювальну складність, м'яка сегментація використовується лише для еталонного локального вікна.

Для обчислення адаптивних ваг  $W_{rf}$  для пікселів еталонного локального вікна, використовується наступна формула:

$$W_{rf}(p, p') = e^{-\frac{|p-p'|}{b_d} - \frac{|ILM_{rf}(p) - ILM_{rf}(p')|}{b_c}}. \quad (2.5)$$

Отримані ваги використовуються для обчислення значення середньозваженої абсолютної різниці інтенсивностей (розмір вікна зіставлення становить  $5 \times 5$ ):

$$WD(p, cnd) = \frac{\sum_{p' \in LW_p} |ILM_{tg}(p + cnd) - ILM_{rf}(p)| \times W_{rf}(p, p')}{W_{rf}(p, p')}, \quad (2.6)$$

де  $p'$  – піксель в локальному вікні  $LW_p$ , центрований відповідно пікселя  $p$  в еталонному зображенні;  $b_d$  та  $b_c$  – пропускну здатність для ядра діапазону та ядра відстані відповідно.

Для запобігання генерації неправильних границь глибини, запропонований метод використовує комплексну міру узгодження, що складається із значення середньозваженої абсолютної різниці інтенсивностей та середньозваженого значення абсолютної різниці градієнта:

$$WDG_h(p, cnd) = \frac{\sum_{p' \in LW_p} |\nabla_h^{tg}(p + cnd) - \nabla_h^{rf}(p)| \times W_{rf}(p, p')}{W_{rf}(p, p')}, \quad (2.7)$$

$$WDG_v(p, cnd) = \frac{\sum_{p' \in LW_p} |\nabla_v^{tg}(p + cnd) - \nabla_v^{rf}(p)| \times W_{rf}(p, p')}{W_{rf}(p, p')}, \quad (2.8)$$

де  $\nabla_h$  та  $\nabla_v$  – горизонтальний та вертикальний градієнт інтенсивності зображення відповідно.

Крім того, запропонований метод адаптивно контролює вплив значення середньозваженої абсолютної різниці інтенсивностей та середньозваженого

значення абсолютної різниці градієнта на обчислення вартості узгодження. Оскільки значення середньозваженої абсолютної різниці інтенсивностей не є робастним, коли різниця в освітленості двох зображень (тобто, еталонного і цільового) є значною, вага середньозваженої абсолютної різниці інтенсивностей в запропонованому методі обернено пропорційна середньому значенню різниці в освітленості:

$$\Omega = \frac{1}{1 + w \times mndif}, \quad (2.9)$$

де  $mndif$  – середнє значення різниці освітленості між зареєстрованими зображеннями;  $w$  – вага  $mndif$ .

Покадрова оцінка послідовності карт глибини призводить до можливого формування часової неузгодженості (часова узгодженість карти глибини означає підтримку плавної та узгодженої інформації про глибину в серії кадрів у відеопослідовності або на декількох часових інтервалах).

Оскільки значення глибини фону не повинні змінюватися між двома сусідніми кадрами, необхідно додати до агрегації витрат член часової узгодженості  $tc$  [69]. Запропонований метод обчислює середню абсолютну різницю для кожного блоку між сусідніми кадрами і визначає, чи є блок фоном відповідно до порогового значення. Член часової узгодженості обчислюється за формулою:

$$tc(p, ds^t, ds^{t-1}) = \omega^t |ds^t - ds^{t-1}|, \quad (2.10)$$

$$\omega^t = \begin{cases} 1, & \text{якщо } p \text{ належить фону} \\ 0, & \text{у іншому випадку} \end{cases}, \quad (2.11)$$

де  $ds^t$  та  $ds^{t-1}$  представляють різницю певного пікселя для поточного та попереднього кадрів відповідно.

Таким чином, остаточний обсяг витрат для «гладких регіонів обчислюється за наступною формулою:

$$C(k, ds^t) = SUM(k, ds^t) + \sum_{p \in r_k} tc(p, ds^t, ds^{t-1}), \quad (2.12)$$

$$C(p, ds^t) = \Omega \times WM(p, ds^t) + (1 - \Omega) \times WDG(p, ds) + tc(p, ds^t, ds^{t-1}). \quad (2.13)$$

*Уточнення обсягу витрат.* Хоча сегментація з перекриттям дозволяє зробити припущення про більшу вірогідність того, що пікселі в кожному сегменті мають подібні диспропорції, велика гладка область може бути розділена на кілька менших областей. У такій ситуації задача зіставлення для безтекстурних областей не може бути розв'язана. Обсяг витрат для «гладких областей» та пікселі в негладких областях є недостовірними, якщо вони не задовольняють критерію унікальності:

$$\frac{C_{ov}^2 - C_{ov}^1}{C_{ov}^1} > T, \quad (2.14)$$

де  $C_{ov}^1$  та  $C_{ov}^2$  – оптимальне значення та субоптимальне значення в обсягу витрат.

Щоб запобігти невідповідності для всієї гладкої області, запропонований метод повторно зіставляє ненадійні гладкі області із зображеннями на базі афінної інваріантної функції за допомогою нормалізованої перехресної кореляції замість суми абсолютних різниць [70]. Якщо повторне зіставлення задовольняє критерію унікальності, витрати на переобчислення в діапазоні розбіжностей для ненадійної згладженої області оновлюються до співрозташування в обсязі витрат. І цей ненадійний згладжений регіон знову стає надійним регіоном.

*Генерація початкової карти глибини.* Після побудови обсягу витрат, для локального вибору найбільшої відмінності  $ds_p$  для певного пікселя  $p$  для негладких областей та  $ds_{sp}$  для супер-пікселя  $sp$  для гладких областей, використовується метод «Winner take all»:

$$d_{sp} = \begin{cases} \arg \max_{ds \in [ds_{min}, ds_{max}]} C(sp, ds^t), & \text{якщо } C(sp, ds^t) \text{ оновлено} \\ \arg \min_{ds \in [ds_{min}, ds_{max}]} C(sp, ds^t), & \text{в іншому випадку} \end{cases}, \quad (2.15)$$

$$ds_p = \arg \min_{ds \in [ds_{min}, ds_{max}]} C(p, ds^t). \quad (2.16)$$

### 2.1.3 Уточнення карти глибини

Оскільки в початкових картах глибин залишилися помилки, пропонується спеціальне уточнення для гладких та негладких регіонів відповідно. Схема виконання процедури уточнення карти глибини зображена на рисунку 2.3.

*Уточнення для гладких регіонів.* Завдяки сегментації з перекриттям, сусідні згладжені області можуть мати подібні достовірні невідповідності. Виходячи з цього припущення, ненадійні згладжені області можна уточнити за допомогою сусідніх надійних згладжених областей.

Запропонований метод виконує пошук в межах сусідів кожної ненадійної згладженої області, а всі надійні згладжені області заносить до списку кандидатів. Якщо середнє значення колірних каналів RGB області  $m(r_i)$  у списку сильно відрізняється від поточної ненадійної згладженої області  $m(r_{current})$ , то область  $m(r_i)$  буде видалено зі списку.

Потім запропонований метод оновлює глибину поточної гладкої області, використовуючи глибину області у списку, колір якої є найближчим до неї. Якщо у списку немає достатньої кількості кандидатів, значення глибини поточної області не змінюється.

Оскільки невідповідності пікселів у кожній згладженій області схожі, але не повністю узгоджуються між собою, необхідно точно налаштувати невідповідність для кожного пікселя в «гладких областях». Для кожного пікселя в «гладких областях» визначається вузький діапазон пошуку на основі поточних невідповідностей у еталонному зображенні.

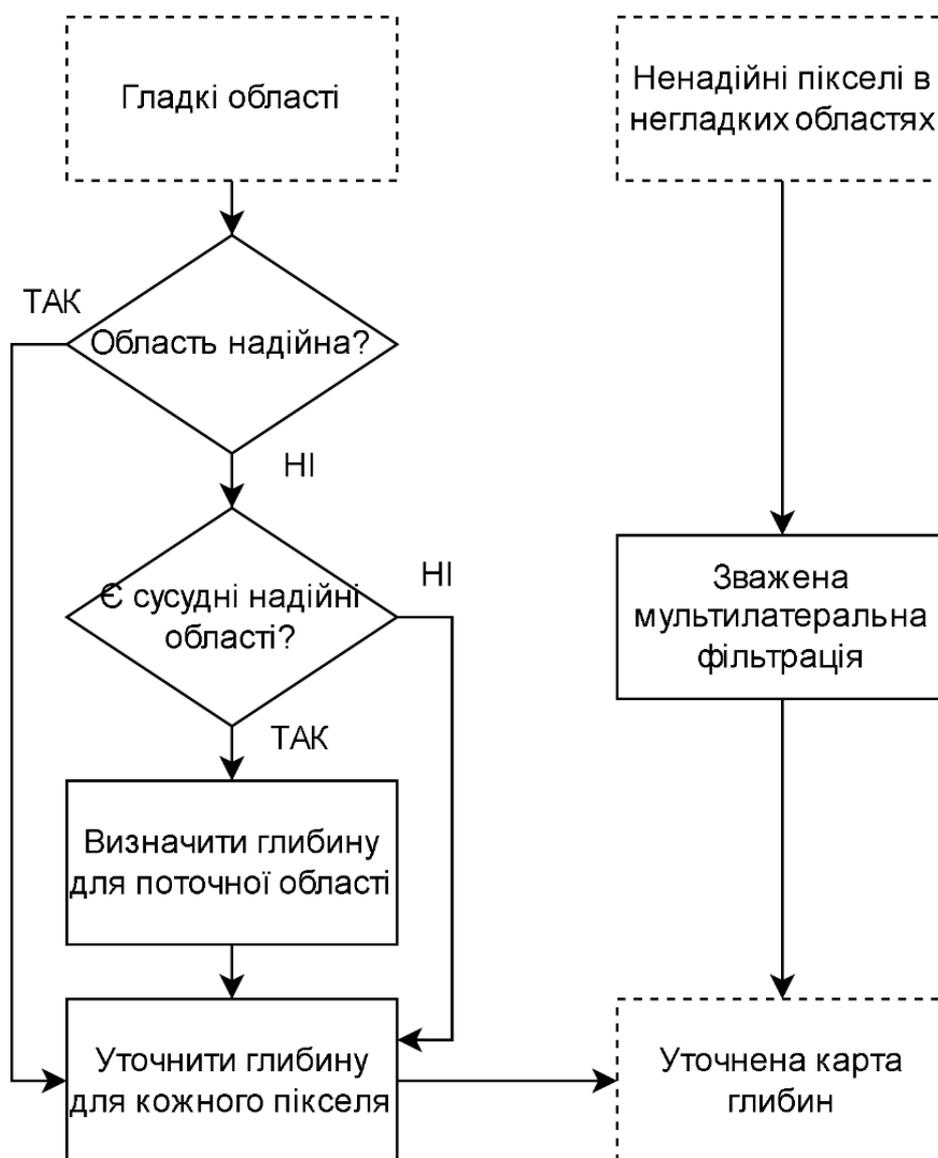


Рисунок 2.3 – Схема виконання процедури уточнення карти глибин

Хоча міра зіставлення на основі кольору погано працює в безтекстурних областях, ознаки на базі афінної інваріантної функції можуть розрізняти пікселі, кольори яких схожі в безтекстурних областях. Виходячи з цього, точне налаштування виконується на зображеннях на базі афінної інваріантної функції на основі міри узгодження нормалізованої перехресної кореляції з локальним вікном  $5 \times 5$ .

У запропонованому методі діапазон розбіжностей обраний як  $d \in [d_{rf} - 2, d_{rf} + 2]$ . Для забезпечення робастності, значення глибини відповідного пікселя

оновлюється лише тоді, коли максимальне значення перевищує 0,85 і задовольняє рівнянню критерію унікальності.

*Уточнення для негладких регіонів.* Для ненадійних пікселів у негладких областях значення їхньої глибини уточнюються за допомогою мультилатеральної фільтрації [71]. У процесі фільтрації ваги пікселів, які належать до ненадійних згладжених областей та ненадійних пікселів у згладжених областях, прирівнюються до нуля. Оскільки обробка виконується лише для ненадійних пікселів, обчислювальна складність суттєво не змінюється.

#### **2.1.4 Експериментальне дослідження запропонованого методу**

Експериментальне дослідження запропонованого методу були проведено на трьох відеопослідовностях – «Akko» [72], «Lovebird2» та «Book arrival» [73] з роздільною здатністю  $640 \times 480$ ,  $720 \times 576$  та  $1024 \times 768$  відповідно. Усі тестові послідовності містять по 100 кадрів.

Робота запропонованого методу порівнювалася з методом DERS без використання механізму розрізів графів [36].

Для об'єктивної оцінки початкових карт глибини, отриманих за допомогою запропонованого методу, в таблиці 2.1 наведено середнє пікове відношення сигнал/шум (PSNR) [74] відрендерених віртуальних зображень, отриманих за допомогою запропонованого методу (з та без використання афінної інваріантної функції), а також за допомогою методу DERS. З таблиці видно, що запропонований метод забезпечує найвищий середній показник PSNR для послідовностей «Akko» та «Lovebird2».

Візуальний аналіз отриманих карт глибин показав, що карти глибини, отримані за допомогою запропонованого методу, виглядають більш наближеними до сцени. Результати методу DERS є досить поганими в безтекстурних областях. Використання афінної інваріантної функції дозволяє виправити деякі ненадійні гладкі ділянки (рис. 2.4).

Таблиця 2.1 – Кількісна оцінка середнього пікового відношення сигнал/шум для тестових послідовностей

	Метод	Метод без АІФ	DERS
Akko	31,79	31,45	31,30
Lovebird2	31,41	31,29	30,89
Book arrival	34,79	34,49	34,89

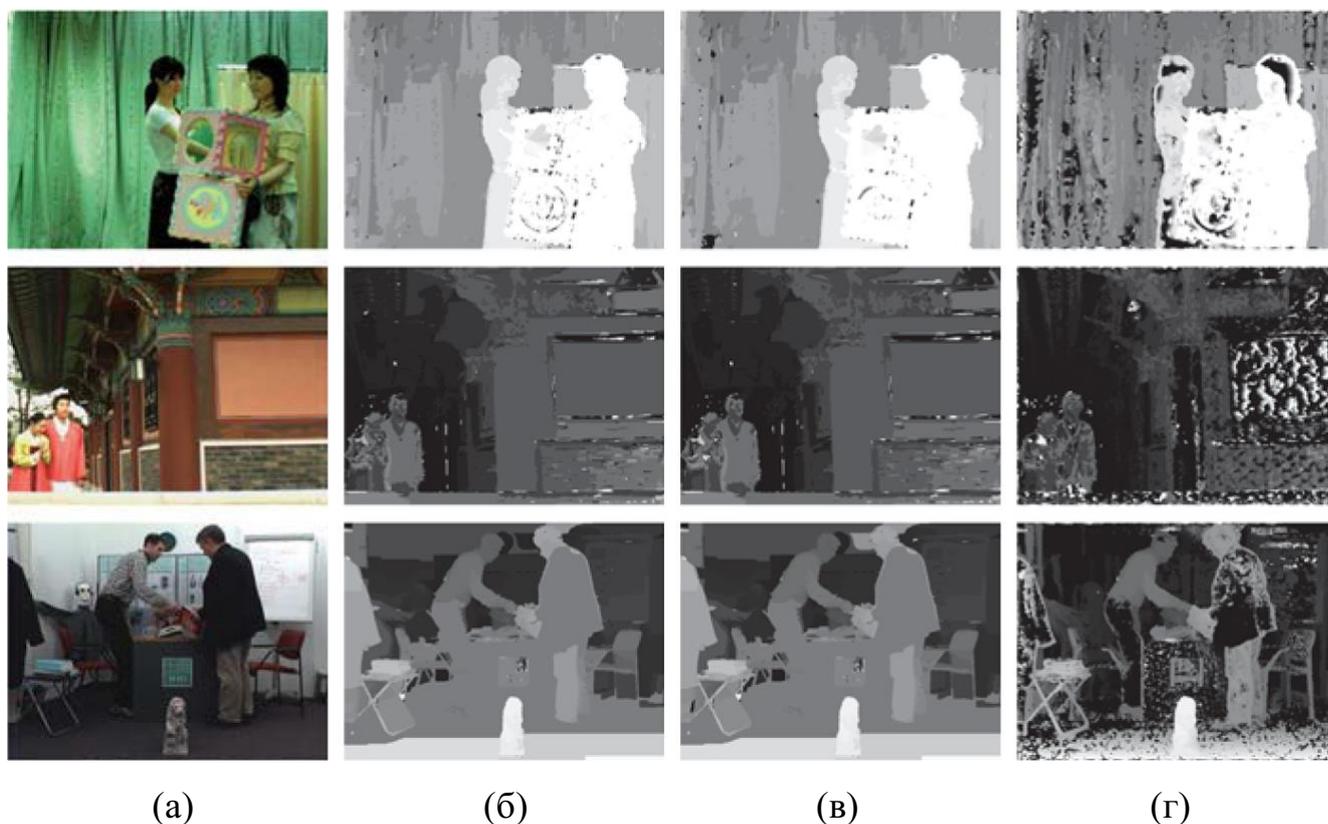


Рисунок 2.4 – Суб'єктивна оцінка початкових результатів: (а) кольорові зображення, карта глибини, отримана за допомогою (б) запропонованого методу, (в) запропонованого методу без афінно-інваріантних ознак, (г) методу DERS

Крім того, глибина області класної дошки в другому рядку, отримана за допомогою запропонованого методу, в основному відповідає дійсності. Але точність глибини в цій області, оцінена за допомогою методу DERS, є відносно нижчою.

Якщо поєднати об'єктивну оцінку з суб'єктивною для «Акко» та «Lovebird2», то можна помітити, що результати запропонованого методу дають найкращі показники. Для послідовності «Book arrival», хоча запропонований метод отримує найкращий результат на більшій частині переднього плану, значення глибини фону оцінюються невірно, особливо на підлозі в нижній частині сцени.

Порівнюючи з результатами запропонованого методу без використання афінно-інваріантних ознак, видно, що використання афінно-інваріантних ознак дає кращі результати, особливо біля голови правого чоловіка.

Причиною неправильної оцінки глибини фону для запропонованого методу є те, що сцена має нахил. У цій ситуації, хоча колір в області підлоги схожий, невідповідності пікселів у цих областях відрізняється один від одного вздовж похилої площини. Тому зіставлення області підлоги є ненадійним.

*Дослідження механізму уточнення карти глибин.* В таблиці 2.2 наведені показники середнього пікового відношення сигнал/шум з використанням механізму уточнення карти глибин та без нього. Як і у випадку дослідження побудови початкової карти глибин, використання механізму уточнення суттєво не відобразилося на кінцевому результаті.

Таблиця 2.2 – Кількісна оцінка середнього пікового відношення сигнал/шум для тестових послідовностей для різних етапів запропонованого методу

	Початкова карта глибин	Уточнена карта глибин
Akko	31,79	31,94
Lovebird2	31,41	31,57
Book arrival	34,79	34,88

*Аналіз обчислювальної складності.* Для всіх методів попиксельного оцінювання їхню обчислювальну складність можна позначити як  $O(TR_dR_w)$ , де  $R_d$  і  $R_w$  – діапазон глибин і фіксований розмір вікна зіставлення, а  $T$  – роздільна здатність кольорового зображення.

На відміну від попiксельних методiв, запропонований метод використовує адаптивну схему зiставлення. Всi пiкселi в однiй гладкiй областi отримують однакову еталонну глибину, а потiм для уточнення глибини кожного пiкселя приймається невеликий дiапазон глибини. В областях з насиченою текстурою для отримання задовiльних результатiв достатньо невеликого розмiру вiкна  $S'_w < S_w$ .

Виходячи з наведеного вище аналізу, обчислювальну складність запропонованого методу можна позначити як  $O(TS'_d S'_w)$ , де  $S'_d < S_d$  – середня дальність пошуку для всiх пiкселiв.

Крiм того, оскiльки кiлькiсть ненадiйних областей i пiкселiв невелика порiвняно з початковою роздiльною здатнiстю зображень, процедура уточнення карти глибин є незначною з точки зору обчислювальної складності. Фактично, бiльша частина часу роботи програми витрачається на сегментацiю еталонного зображення.

Загальний середнiй час роботи та час роботи всiх етапiв методу, включаючи сегментацiю, агрегування вартостi та уточнення, наведено в таблицi 2.3.

Таблиця 2.3 – Порiвняння середнього часу роботи запропонованого методу та методу-аналогу

	DERS	Запропонований метод			
		Сегментацiя	Агрегування вартостi	Уточнення	Всього
Akko	8,3 сек.	4,8 сек.	1,4 сек.	0,096 сек.	6,3 сек.
Lovebird2	9,7 сек.	6 сек.	1,7 сек.	0,14 сек.	7,8 сек.
Book arrival	28,7 сек.	18 сек.	4,3 сек.	0,7 сек.	23 сек.

## 2.2 Покращення часової узгодженостi для послiдовностей карт глибини

У багатьох системах технiчного зору, для аналізу змiни сцени відбувається покадрове обчислення та аналіз карт глибини, що дозволяє побудувати вiртуальне представлення сцени в часi. Під час побудови таких послiдовностей дуже часто з'являються артефакти миготiння iз-за того, що значення глибини певних областей,

які мали б мати однакову глибину на сусідніх кадрах, відрізняються з очевидними розривами. Проблема появи артефактів на послідовності карт глибини часто називають проблемою часової узгодженості. Проблема часової узгодженості також знижує ефективність стиснення послідовності карт глибини, тому що доводиться виділяти більше бітів для збереження такої штучної неузгодженості.

В рамках даного дисертаційного дослідження пропонується метод покращення часової узгодженості для багатоканальних послідовностей карт глибини.

Хоча багато методів побудови карт глибини (наприклад, DERS) виконують оцінку шляхом створення темпорального члена в енергетичній функції, такий член важко збалансувати з іншими членами між кадрами, що призводить до нестабільних результатів. В рамках роботи [1] був запропонований метод усереднення статичної сцени сусідніх кадрів для покращення часової узгодженості. Однак він не враховує надійність значень глибини, які беруть участь в усередненні, що сприяє поширенню помилок.

На відміну від інших методів, запропонований метод покращення часової узгодженості враховує надійність показника глибини та використання динамічних атрибутів регіонів.

### **2.2.1 Принцип роботи запропонованого методу**

На рисунку 2.5 проілюстрований принцип роботи запропонованого методу покращення часової узгодженості. Запропонована адаптивна вага у середньозваженій вазі складається з двох частин:

- 1) ймовірність належності до статичної сцени;
- 2) надійність значень глибини.

Враховуючи ймовірність належності пікселів до статичної сцени, можна зменшити поширення помилок на рухомі області. А надійність значень глибини може зменшити вплив недостовірної глибини в середньозваженому значенні, що зменшує розповсюдження помилки між сусідніми кадрами.



Рисунок 2.5 – Принцип роботи запропонованого методу

У той же час, оскільки передбачається, що послідовність може бути змодельована як Марківська модель, попередній і наступний кадри мають найсильніший зв'язок з кадром, який потрібно покращити. Тому запропоноване середньозважене значення обмежується двома сусідніми кадрами та описується наступною формулою:

$$DT'_c = \frac{v^p \alpha^p DT'_p + \alpha^c DT_c + v^n \alpha^n DT_n}{v^p \alpha^p + \alpha^c + v^n \alpha^n}, \quad (2.17)$$

де  $DT_c$ ,  $DT_p$  та  $DT_n$  – значення глибини поточного, попереднього і наступного кадрів послідовності карт глибини відповідно,  $DT'$  – значення глибин покращеного кадру послідовності карт глибини,  $v^p$  та  $v^n$  – ймовірність належності до статичної сцени для всіх пікселів у попередньому та наступному кадрах послідовності,  $\alpha^p$ ,  $\alpha^c$  та  $\alpha^n$  – достовірність значень глибини для всіх пікселів у попередньому, поточному та наступному кадрах глибини відповідно.

Для спрощення рівняння всі значення глибини  $dt$  у певному кадрі глибини, що підлягає покращенню, зібрані в  $DT$ , а в рівнянні (2.17) всі пікселі обробляються поелементно, а не матричним множенням.

*Проведемо обчислення ймовірності належності до статистичної сцени.*

На основі спостереження, що статичні області мають однаковий колір між сусідніми кадрами, для оцінки належності пікселів до статичних областей використовується метод середнього абсолютного відхилення (MAD). Оцінювання виконується у локальному вікні з центром у пікселі, який оцінюється. На основі двостороннього фільтра, який представлений в роботі [76], надійність статичності для пікселя  $k$  моделюється наступним чином (для надійності така вага обрізається до 0 для пікселів, що належать до рухомих областей з великим MAD):

$$v(k) = \begin{cases} e^{\frac{-MAD_v(r_v^k, t_v^k)}{\delta_v}}, & MAD(r_v^k, t_v^k) < T_s \\ 0 & , MAD(r_v^k, t_v^k) \geq T_s \end{cases} \quad (2.18)$$

де  $v(k)$  – елемент  $v$ , що представляє вагу  $k$ -го пікселя в кадрі глибини.  $r_v^k$  та  $t_v^k$  - колір пікселів у локальних вікнах з центром в  $k$  в двох сусідніх кадрах глибини, Таким чином,  $v^p$  обчислюється між поточним і попереднім кадрами глибини, а  $v^n$  обчислюється між поточним і наступним кадрами глибини. Параметри  $T_s$  та  $\delta_v$  є заздалегідь визначеними значеннями.

*Проведемо обчислення надійності значень глибини.*

Через наявність шуму в картах глибин, при їх обчисленні слід явно враховувати показник надійності значень глибини, що дає більшу вагу значенню глибини з більшим показником надійності. Для кожного пікселя на карті глибини поточного ракурсу, оскільки відповідний піксель в іншому ракурсі можна обчислити за допомогою дисперсії, надійність глибини можна оцінити за допомогою середнього абсолютного відхилення між локальними вікнами, центрованими на такій парі пікселів.

Щоб врахувати оклюзію, виконується обчислення для лівого та правого сусідніх ракурсів і вибирається менше середнє абсолютне відхилення. Подібно до обчислення ймовірності належності до статичної сцени, надійність показника глибини для пікселя  $k$  також моделюється як експоненціальна функція:

$$\alpha(k) = e^{\frac{-MAD^k}{\delta_\alpha}} \quad (2.19)$$

$$MAD^k = \min\left(MAD(r_\alpha^k, lt_\alpha^{k+d_k}), MAD(r_\alpha^k, rt_\alpha^{k-d_k})\right) \quad (2.20)$$

де  $\alpha(k)$  – елемент  $\alpha$ , що представляє достовірність значення  $k$ -го пікселя у кадрі глибини,  $d_k$  – значення глибини для  $k$ . Таким чином, відповідні пікселі для  $k$  у поточному ракурсі мають значення  $k + d_k$  та  $k - d_k$  у лівому та правому ракурсах відповідно.  $r_\alpha^k$  – локальне вікно з центром у  $k$  у поточному ракурсі,  $lt$  та  $rt$  – карти глибини у лівому та правому суміжних ракурсах відповідно,  $\delta_\alpha$  – заздалегідь визначений параметр.

*Проведемо обчислення вибору розмір улокального вікна*

Оскільки ймовірність належності до статичної сцени та достовірність значень глибини обчислюються в межах локального вікна, необхідно адаптивно підбирати розмір таких вікон.

Для обчислення достовірності глибини потрібно, щоб у локальному вікні було якомога більше пікселів з однаковою глибиною, тому перевага надається малому розміру вікна. Однак, оцінка у великому вікні є більш стійкою до шуму. У запропонованому методі розроблено збалансовану схему, яка визначає множину розмірів вікон, зокрема  $9 \times 9$ ,  $7 \times 7$ ,  $5 \times 5$  та  $3 \times 3$ .

Спочатку вибирається найбільший розмір і обчислюється дисперсія для такого локального вікна на карті глибини. Якщо значення дисперсії більше заданого порогу, це означає, що значення глибин в межах локального вікна не є подібними. Тому слід зменшити розмір вікна і обчислити дисперсію у новому вікні. Робота алгоритму зупиняється, якщо значення дисперсії стане меншим за поріг або буде обрано найменший розмір вікна.

## 2.2.2 Експериментальне дослідження роботи методу

В процесі експериментального дослідження роботи методу покращення часової узгодженості, були проведені наступні заходи:

- 1) оцінка часової узгодженості безпосередньо на покращених картах глибини;
- 2) оцінка показника PSNR на рендерованих кольорових зображеннях;
- 3) оцінка ефективності кодування покращених карт глибини;
- 4) суб'єктивна оцінка часової узгодженості.

Дослідження запропонованого методу було проведено на послідовностях «Akko», «Lovebird2» та «Newspaper» [77], кожна з яких має п'ять ракурсів та 100 кадрів для кожного ракурсу. Початкові послідовності карт глибин були отримані за допомогою методів DERS [57], Fu [75] та за допомогою запропонованого методу.

*Оцінка часової узгодженості безпосередньо на покращених картах глибини.*

Часова узгодженість оцінюється на позначених ділянках, показаних на рисунку 2.6, які є статичними сценами. У таких ділянках глибина пікселів на одному і тому ж місці на всіх кадрах збирається як випадкова величина. Дисперсія кожної змінної обчислюється незалежно. Середнє значення всіх дисперсій  $\psi$  використовується як метрика для обчислення.

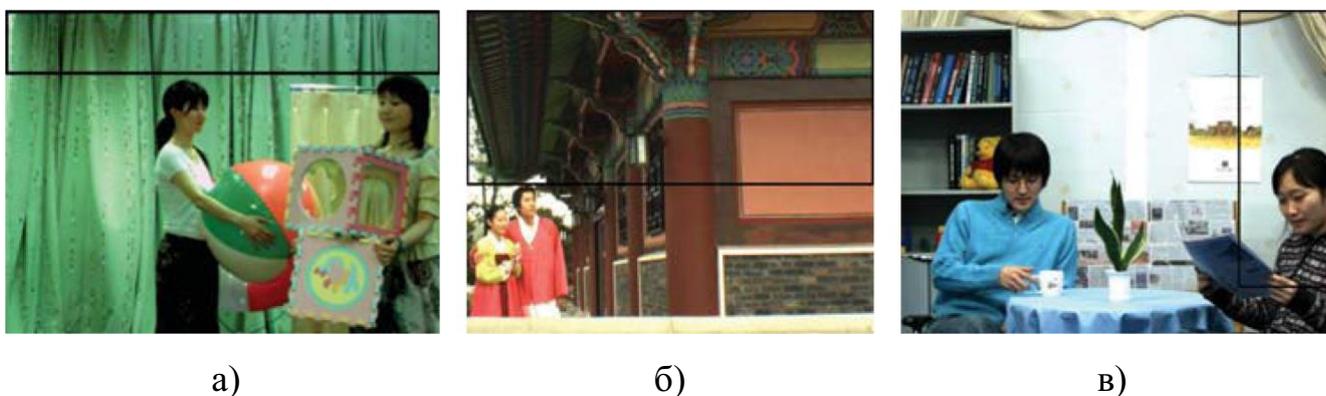


Рис. 2.6 – Зображення статичних регіонів для тестових послідовностей (а) Akko, (б) Lovebird2, (в) Newspaper

В таблиці 2.4 наведені результати роботи методів покращення часової узгодженості для двох ракурсів. Результати експериментального дослідження показують, що запропонований метод має найменше значення  $\psi$ , що доводить його ефективність.

Таблиця 2.4 – Оцінка часової узгодженості для статичних регіонів

Послідовність	Метод		
	DERS	Fu	Запропонований
Akko	Ракурс 27: 4,13779 Ракурс 29: 4,23298	Ракурс 27: 2,97221 Ракурс 29: 2,88042	Ракурс 27: 2,16257 Ракурс 29: 2,18862
Lovebird2	Ракурс 8: 5,74498 Ракурс 10: 1,02449	Ракурс 8: 4,10580 Ракурс 10: 0,32439	Ракурс 8: 3,29369 Ракурс 10: 0,22409
Newspaper	Ракурс 4: 2,68077 Ракурс 6: 3,10668	Ракурс 4: 1,50019 Ракурс 6: 1,8968	Ракурс 4: 1,36010 Ракурс 6: 1,588512

#### *Оцінка показника PSNR на рендерованих кольорових зображеннях*

Для подальшої валідації запропонованого методу оцінено якість рендерованих кольорових зображень з використанням покращених карт глибини. В таблиці 2.5 наведено середні значення PSNR для рендерованих кольорових зображень. Результати експериментального дослідження показують покращення карт глибини при використанні запропонованого методу.

Таблиця 2.5 – Порівняння середнього PSNR рендерованих кольорових зображень

Послідовність	Метод		
	DERS	Fu	Запропонований
Akko	33.69 dB	33.72 dB	33.77 dB
Lovebird2	31.64 dB	31.68 dB	31.74 dB
Newspaper	32.68 dB	32.70 dB	32.75 dB

#### *Оцінка ефективності кодування покращених карт глибини.*

В даному експерименті проводилася оцінка отриманого бітрейту відео послідовності карт глибини, а також середнього значення PSNR рендерованих кольорових зображень з використанням відповідних декодованих карт глибини.

Для тестування були використані послідовності «Akko», «Lovebird2» та «Newspaper» з трьома значеннями точок квантування (QP) – 22, 27 та 32 для кожної послідовності. В таблиці 2.6 наведені результати експериментального дослідження роботи запропонованого методу у порівнянні з методом DERS. Результати дослідження показують, що послідовність карт глибини, покращена запропонованим методом, не тільки значно знижує бітрейт, але й покращує якість відтворення зображень.

Таблиця 2.6 – Порівняння продуктивності кодування послідовності карт глибини

Послідовність	QP	Метод					
		BR-DERS, kbit/s	PSNR-DERS, dB	BR-власний, kbit/s	PSNR-власний, dB	$\nabla$ BR, %	$\nabla$ PSNR, dB
Akko	22	759	34,29	670	34,48	-13,3	+0,19
	27	430	34,35	339	34,51	-26,8	+0,06
	32	229	34,34	177	34,47	-28,4	+0,13
Lovebird2	22	1560	31,59	1139	31,71	-37,0	+0,12
	27	860	31,65	587	31,74	-46,5	+0,09
	32	435	31,69	287	31,81	-51,6	+0,12
Newspaper	22	430	32,59	356	32,62	-20,7	+0,03
	27	230	32,64	189	32,68	-21,7	+0,04
	32	119	32,66	103	32,67	-15,5	+0,01

### *Суб'єктивна оцінка часової узгодженості.*

Оскільки запропонований метод покращує темпоральну узгодженість для статичних областей, щоб зменшити артефакти миготіння на зображенні, середній PSNR не отримує значного покращення.

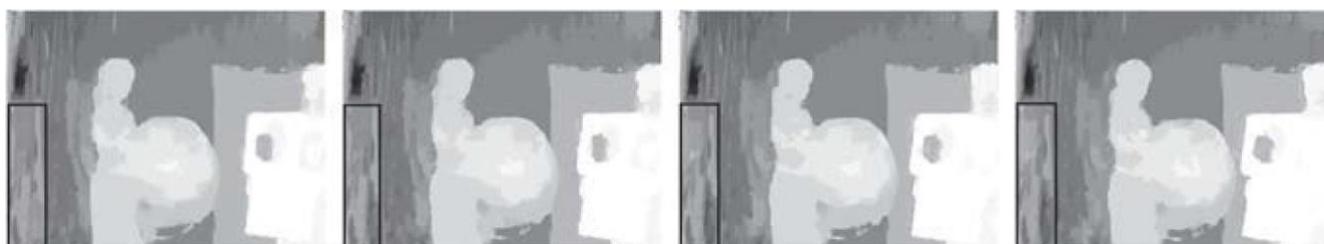
Однак покращення часової узгодженості можна побачити за допомогою суб'єктивного оцінювання, як показано на рисунку 2.7, на якому наведені декілька сусідніх карт глибини.



а)



б)



в)

Рисунок 2.7 – Суб'єктивна оцінка часової узгодженості, (а) 4 послідовні кадри кольорової послідовності (б) відповідні 4 послідовні карти глибини, покращені за допомогою методу DERS (в) відповідні 4 послідовні карти глибини, покращені за допомогою запропонованого методу

### 2.3 Метод наближеного стереозіставлення на основі алгоритму StereoBM з використанням вейвлет-перетворення Гаара

У контексті побудови карти глибин на основі стереозображень ключовим етапом є точна і ефективна оцінка диспаритету між відповідними пікселями лівого та правого зображення.

Традиційні підходи, такі як методи на основі блочного порівняння (наприклад, SAD або SSD), демонструють обмежену точність у випадках з

малоконтрастними або безтекстурними областями, а також мають значне енергоспоживання при апаратній реалізації.

З метою підвищення ефективності та зменшення обчислювальних витрат у роботі запропоновано наближений метод побудови карти глибин із використанням простору вейвлет-перетворення з фокусом на адаптивному виборі довжини носія вейвлет-функції Гаара.

### 2.3.1 Формальна постановка задачі та суть методу

Нехай задано пару каліброваних зображень  $I_L(x, y)$  та  $I_R(x, y)$ , отриманих з лівої та правої камер стереосистеми відповідно.

Завдання полягає у побудові карти диспаратетів  $D(x, y)$ , яка наближено описує глибину сцени. Задача формалізується як мінімізація функції вартості суміщення  $C(x, y, d)$ :

$$D(x, y) = \arg \min_{d \in D} C(x, y, d), \quad (2.21)$$

де  $D$  – допустимий інтервал зміщень (диспаратетів);

$C(x, y, d)$  – функція вартості, яка визначає міру подібності між вікном  $W_L \text{ in } \subset I_L$  та відповідним зсунутих вікном  $W_R(d) \subset I_R$ .

**Суть методу.** Запропонований метод базується на аналізі структурних особливостей в локальних ділянках зображення за допомогою вейвлет-перетворення Гаара. Відмінною рисою методу є здійснення попередньої локалізації екстремуму у просторі коефіцієнтів вейвлет-перетворення, що дозволяє звужити область пошуку та значно зменшити обсяг обчислень.

Після локалізації потенційної області збігу виконується поступове уточнення позиції екстремуму шляхом варіювання довжини носія вейвлет-функції Гаара. Загальна схема методу включає такі кроки:

*Крок 1. Попередня обробка вхідних зображень.*

Стереопара (ліва та права проєкції сцени) конвертується у відтінки сірого, після чого застосовується згладжування для зменшення шуму (наприклад, за допомогою гаусового фільтру). Це покращує стабільність відповідності в подальших етапах.

*Крок 2. Вейвлет-перетворення зображень.*

До кожного зображення застосовується дискретне вейвлет-перетворення Гаара (DWT), що дозволяє розкласти сигнал на чотири піддіапазони:  $LL$ ,  $LH$ ,  $HL$ ,  $HH$ . Найбільшу інформативність у контексті локалізації глибини мають компоненти  $HL$  та  $LH$ , які відповідають горизонтальним і вертикальним градієнтам. Формально, перетворення описується як:

$$W(x, y) = \sum_{i,j} I(i, j) \cdot \psi_{x,y}(i, j), \quad (2.22)$$

де  $\psi_{x,y}$  – вейвлет Гаара;

$I(i, j)$  – піксель зображення.

*Крок 3. Вибір фрагментів для стереозіставлення.*

На основі коефіцієнтів вейвлет-перетворення обираються найбільш контрастні області (екстремальні значення локальної енергії коефіцієнтів), що зменшує кількість точок для пошуку відповідності без втрати інформативності.

*Крок 4. Побудова карти диспаритету (disparity map).*

На вибраних регіонах виконується модифіковане стереозіставлення за алгоритмом StereoBM. Відстань між пікселями обчислюється з використанням метрики SAD (сума абсолютних різниць) або комбінації SAD + градієнтних різниць (SGRAD):

$$C(x, d) = \sum_{(i,j) \in W} |I_L(i, j) - I_R(i - d, j)|, \quad (2.23)$$

де  $d$  – зміщення;

$W$  – вікно.

*Крок 5. Інтерполяція та добудова карти глибин.*

Оскільки StereoBM працює тільки на вибраних зонах, виконується згладжування карти глибин і заповнення прогалів за допомогою білінійної або сплайн-інтерполяції. В окремих випадках застосовується маскуванню невизначених ділянок за допомогою додаткових умов згладженості.

*Крок 6. Постобробка карти глибини.*

Для покращення візуальної та метричної якості карти глибини застосовуються морфологічні операції (наприклад, усунення шумів або сплесків) та фільтрація з використанням медіанного фільтра.

### **2.3.2 Експериментальне дослідження запропонованого методу**

Для експерименту було встановлено такі параметри: початкова область пошуку диспаритету –  $[1; 1]$ , крок дискретизації – 1, довжина носія  $a = 17$ . Застосування ітеративного пошуку екстремуму дозволило зійтись до мінімуму функції за 3 ітерації.

Для кількісної оцінки ефективності модифікованого методу, позначеного як StereoBM-Mod, було проведено серію експериментів на підмножині даних з GIST Stereo Video Dataset [78], зокрема на відеопослідовностях BookArrival (відео 1), Balloons (відео 2), Cafe (відео 3) та Newspaper (відео 4). Ці дані були обрані через наявність реалістичних сцен з неоднорідним освітленням, частковими перекриттями об'єктів і складними текстурями, що дозволяє оцінити стабільність роботи алгоритму в умовах, наближених до реальних.

Розміри кадрів становили  $640 \times 480$  пікселів, а референтні карти глибини були отримані за допомогою структурованого світла з подальшим ручним коригуванням, що забезпечило надійну основу для оцінки точності.

У якості методів аналогів були обрані наступні методи:

– StereoBM (Stereo Block Matching) – базовий метод блочного зіставлення, що порівнює фрагменти лівого та правого зображень за метрикою, наприклад,

SAD. Має високу швидкодію, але низьку точність у слабо текстурованих зонах [79].

– SGBM (Semi-Global Block Matching) – напівглобальний метод, який поєднує локальні оцінки подібності з глобальною регуляризацією. Забезпечує кращу точність на межах об'єктів, але потребує більше обчислень [80].

– ELAS (Efficient Large-scale Stereo Matching) – ефективний масштабований метод, що використовує опорні точки та триангуляцію для побудови карти диспаратностей. Дає добрі результати у складних сценах [81].

– WLS (Weighted Least Squares) – ваговий метод найменших квадратів для згладжування карти диспаратності. Зменшує шум, зберігаючи структурні межі об'єктів. Застосовується як постпроцесинг [82].

Оцінювання виконувалося за трьома ключовими критеріями: точність реконструкції глибини (вимірювана за допомогою метрики PSNR), енергоспоживання та час обробки.

В таблиці 2.7 наведені результати експериментального дослідження запропонованого методу та методів-аналогів. На основі графіка, що відображає значення середнього пікового відношення сигнал/шум (рис. 2.8), встановлено, що запропонований алгоритм StereoBM-Mod забезпечує найвищу якість реконструкції глибини серед усіх протестованих методів.

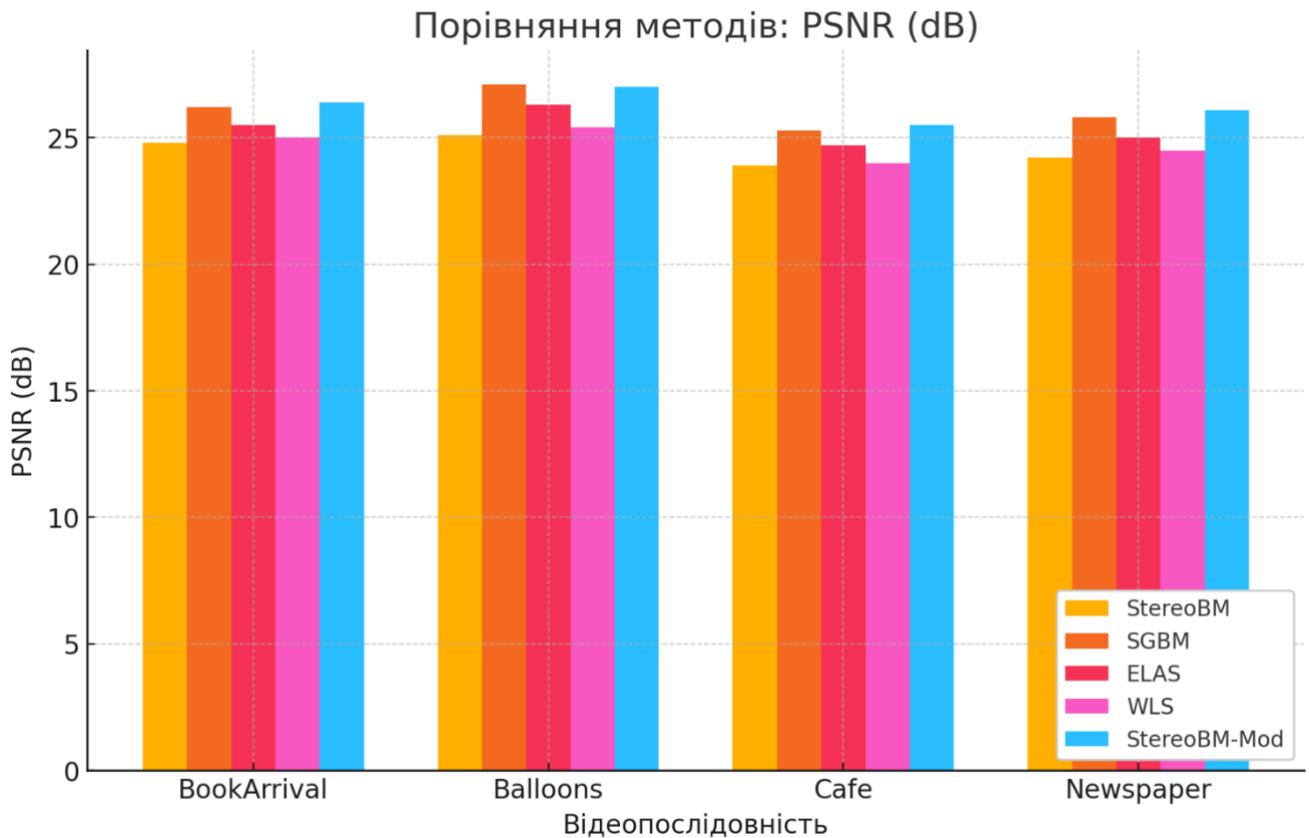


Рисунок 2.8 – Графік порівняння методів за показником точності

Для всіх чотирьох відеопослідовностей спостерігається стабільне перевищення показників PSNR у порівнянні з базовим StereoBM (на 1.5–2.4 дБ), що вказує на суттєве зменшення помилок реконструкції та кращу відповідність отриманих глибинних карт еталонним.

Таблиця 2.7 – Порівняння роботи методів стереозіставлення

Метод	Точність (PSNR, дБ)				Споживання струму, мА				Час обробки, мсек.			
	Відеопослідовність				Відеопослідовність				Відеопослідовність			
	1	2	3	4	1	2	3	4	1	2	3	4
StereoBM	24,8	25,1	23,9	24,2	4,5	4,6	4,4	4,5	308	316	307	311
SGBM	26,2	27,1	25,3	25,8	4,2	4,3	4,1	4,2	282	284	273	283
ELAS	25,5	26,3	24,7	25,0	3,9	4,0	3,8	3,9	265	267	254	265
WLS	25,0	25,4	24,0	24,5	4,0	4,1	3,9	4,0	272	273	268	277
StereoBM-Mod	26,4	27,0	25,5	26,1	3,2	3,3	3,1	3,2	243	244	237	241

Алгоритм SGBM, що використовує напівглобальну оптимізацію, демонструє дещо нижчі, але конкурентоспроможні значення точності, зокрема на послідовностях BookArrival та Cafe. Водночас, методи ELAS та WLS демонструють середні показники точності, що варіюються в межах 24.8–25.5 дБ. Найнижчі значення спостерігаються у базового StereoBM, що є очікуваним з огляду на його просту локальну природу зіставлення, чутливість до шуму та недостатню здатність до врахування структурної інформації сцени. Таким чином, запропонований модифікований алгоритм продемонстрував підвищену точність, що обґрунтовується кращим виявленням структурних елементів сцени за рахунок локального контекстного аналізу, який реалізується через вейвлет-фільтрацію. Це свідчить про доцільність використання модифікованого підходу в умовах обмежених обчислювальних ресурсів при збереженні високої якості реконструкції.

На основі графіка, що демонструє показники споживання електроенергії (рис. 2.9), встановлено, що найбільш енергоефективними є алгоритми StereoBM та StereoBM-Mod.

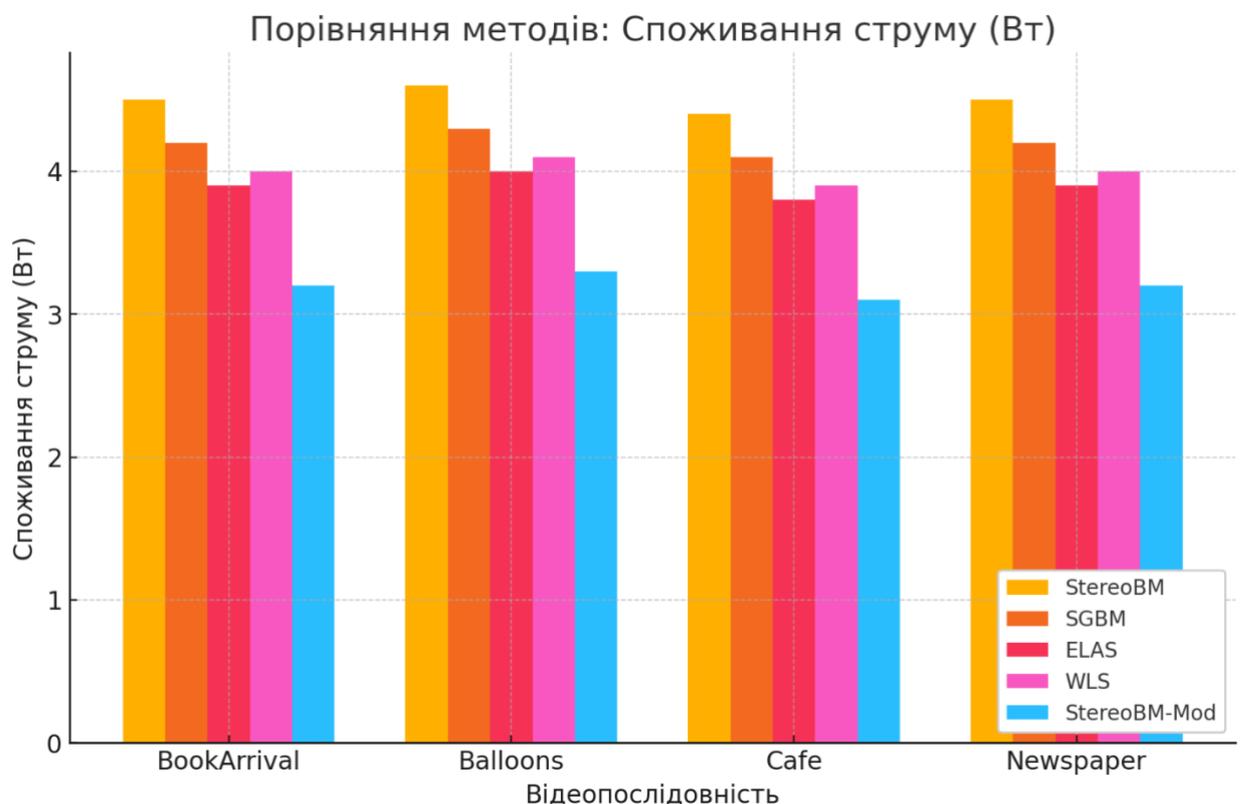


Рисунок 2.9 – Графік порівняння методів за показником споживання енергії

Зокрема, модифікований варіант, незважаючи на додаткову обчислювальну складність, демонструє лише незначне зростання енергоспоживання (у межах 3–6% у порівнянні з базовим StereoBM), що свідчить про доцільність його використання в автономних або вбудованих системах.

Алгоритми SGBM та ELAS, які передбачають складні процедури глобальної оптимізації або семантичного вирівнювання, вимагають суттєво більших енергетичних витрат – до 1.6–1.8 разів більше у порівнянні з StereoBM-Mod. Особливо це помітно на відеопослідовностях Cafe та Newspaper, де ці методи демонструють максимальні пікові значення енергоспоживання. Метод WLS, який використовує фільтрацію із зваженими локальними площинами, має проміжні показники, але все ж поступається за енергоефективністю запропонованому підходу.

Таким чином, модифікований метод зберігає прийнятний баланс між точністю та енергоспоживанням, що є важливою перевагою в контексті мобільних або енергонезалежних платформ.

Третій графік демонструє час обробки зображень у мілісекундах (мс), що є критичним параметром для систем реального часу (рис. 2.10).

Найменші значення часу спостерігаються для алгоритму StereoBM, що відповідає його спрощеній реалізації та відсутності складних оптимізаційних процедур.

Водночас, StereoBM-Mod демонструє лише незначне зростання часу (у межах 10–15%) порівняно з оригіналом, що є цілком прийнятним з урахуванням значного підвищення точності.

Методи SGBM, ELAS та WLS мають значно вищий час обробки, особливо на відеопослідовностях з високою щільністю текстур (BookArrival, Cafe). У випадку ELAS, час перевищує 100 мс на кадр, що фактично унеможливорює його використання в режимі реального часу без апаратного прискорення. Аналогічно, WLS демонструє часові затримки, пов'язані з процедурою розгортання локальних згладжувальних фільтрів.

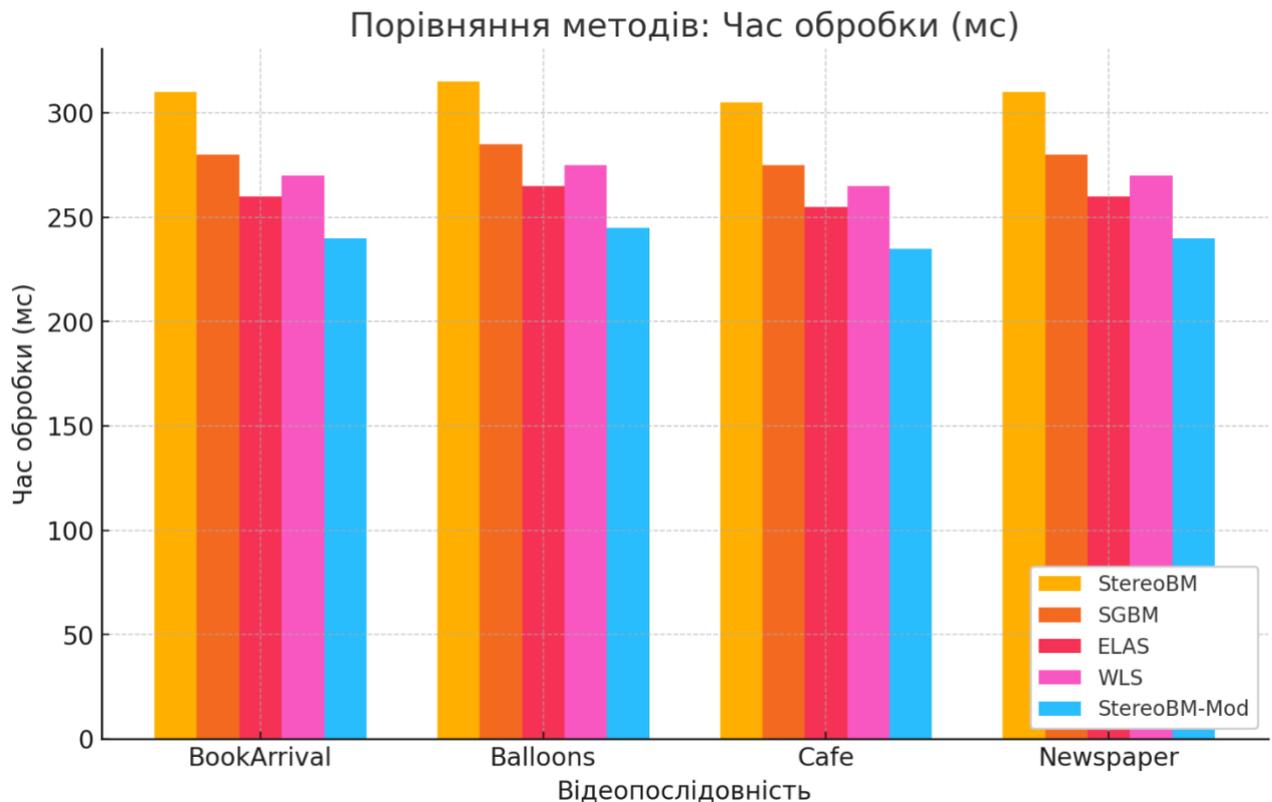


Рисунок 2.10 – Графік порівняння методів за показником часу обробки

Таким чином, StereoBM-Mod є оптимальним компромісом між швидкістю й точністю, зберігаючи продуктивність, достатню для інтеграції у системи реального часу (наприклад, візуальна навігація, робототехніка, контроль доступу). Проведене дослідження демонструє, що запропонований модифікований алгоритм StereoBM-Mod, який базується на класичному StereoBM із додатковим етапом вейвлет-обробки, забезпечує оптимальне поєднання трьох важливих характеристик: високої точності, низького енергоспоживання та малого часу обробки. Це дозволяє рекомендувати даний метод для застосування в обчислювально обмежених середовищах, де важлива як ефективність, так і швидкодія алгоритмічного забезпечення.

### 2.3.3 Обговорення результатів дослідження

У цьому підрозділі було запропоновано та експериментально досліджено модифікований метод побудови карти глибини на основі класичного алгоритму

StereoBM, вдосконаленого шляхом попередньої вейвлет-обробки за допомогою фільтра Гаара. Запропонований підхід поєднує переваги простоти реалізації, низької обчислювальної складності та покращеної точності, що досягається за рахунок ефективного виділення контурних структур на етапі попередньої обробки.

Для оцінки ефективності методу було проведено серію експериментів на реальних відеопослідовностях з GIST-датасету (BookArrival, Balloons, Cafe, Newspaper), а також виконано порівняння з іншими поширеними підходами, зокрема StereoBM, SGBM, ELAS та WLS. Аналіз результатів засвідчив, що:

За точністю (PSNR) StereoBM-Mod стабільно перевершує класичний StereoBM, поступаючись лише складнішим глобальним алгоритмам, водночас демонструючи відчутно нижчу енергетичну й часову вартість.

Метод зберігає низьке енергоспоживання, що є критичним для вбудованих і мобільних систем, поступаючись лише класичному StereoBM, проте переважаючи інші сучасні методи.

За швидкістю модифікований алгоритм залишається придатним для використання в режимі реального часу, демонструючи значну перевагу над більш складними методами (SGBM, ELAS, WLS).

Таким чином, запропонований метод StereoBM-Mod забезпечує збалансовану ефективність за трьома критеріями: точністю, енергоспоживанням та часом обробки. Це робить його доцільним вибором для широкого спектра практичних застосувань у системах технічного зору, зокрема в задачах візуальної навігації, автоматизованого контролю та розпізнавання об'єктів у реальному часі.

## **2.4 Висновки до другого розділу**

В даному розділі дисертаційної роботи запропоновані методи удосконалення пасивних методів побудови карт глибини.

Для підвищення швидкості пасивних методів побудови карти глибини за допомогою множини кольорових зображень був запропонований локальний метод, що базується на адаптивних вагах, процедурі адаптивного зіставлення та афінно-

інваріантних ознаках. Метод включає початкову оцінку глибини шляхом агрегації витрат з використанням адаптивного зіставлення, подальше уточнення витрат у гладких регіонах та побудову карти глибини за принципом «переможець отримує все». Далі виконується етап уточнення карти глибини: для гладких ненадійних регіонів проводиться пошук надійних сусідніх областей з подібним кольором для корекції їхніх значень глибини, а для негладких регіонів застосовується багатостороння фільтрація, під час якої ненадійні пікселі виключаються з розрахунків. Експериментальні дослідження на різних наборах даних показали зменшення часу на побудову карти глибини на 19,9% – 24,1% у порівнянні з методом-аналогом.

Для зменшення кількості артефактів при формуванні віртуальних зображень з послідовностей карт глибини запропоновано метод, що покращує часову узгодженість шляхом введення показника надійності глибини та динамічних атрибутів регіонів. Метод складається з двох етапів: на першому оцінюється належність областей до статичних регіонів за середнім значенням абсолютної різниці кольору сусідніх кадрів у локальному вікні; на другому етапі визначається надійність значень глибини, враховуючи шум та оклюзії, за допомогою порівняння середніх абсолютних різниць у сусідніх видах. Експериментальні дослідження на різних наборах даних показали збільшення часової узгодженості на 10,3% – 44,75% у порівнянні з методом-аналогом та покращення ефективності кодування карти глибини на 13,3% – 51,6% у порівнянні з розглянутим методом-аналогом.

Для підвищення точності та зменшення обчислювальних витрат при побудові карти глибин у системах стереозору, було запропоновано модифікований метод StereoBM-Mod, що базується на класичному алгоритмі StereoBM із попередньою обробкою вейвлет-перетворенням Гаара. Використання вейвлет-аналізу дозволило локалізувати найбільш інформативні ділянки зображення та звузити область пошуку відповідей. Експериментальні результати засвідчили, що запропонований підхід забезпечує стабільне підвищення точності реконструкції глибини (на 1.5–2.4 дБ за PSNR) у порівнянні з базовим StereoBM, зберігаючи при цьому низьке енергоспоживання та придатність для роботи в реальному часі.

Отримано наступні пункти наукової новизни:

– *удосконалено* пасивний локальний метод побудови карти глибини за рахунок використання процедури адаптивного зіставлення та афінно-інваріантних ознак. Використання запропонованого методу дозволило підвищити швидкість побудови карти глибини;

– *удосконалено* метод покращення часової узгодженості для послідовностей карт глибини за рахунок впровадження показника надійності глибини та використання динамічних атрибутів регіонів. Використання запропонованого методу дозволило покращити часову узгодженість та покращити ефективність кодування карти глибини;

– *набув подальший розвиток* метод локального стереозіставлення StereoVM шляхом інтеграції наближеного вейвлет-перетворення Гаара для попереднього виділення контурних ознак та локалізації екстремумів. Використання запропонованого методу дозволило зменшити обчислювальну складність процесу побудови карти глибини, зберігаючи при цьому точність у слабо текстурованих зонах

### 3 МЕТОДИ УДОСКОНАЛЕННЯ КАРТ ГЛИБИНИ, ОТРИМАНИХ ЗА ДОПОМОГОЮ АКТИВНИХ МЕТОДІВ

Удосконалення карт глибини, отриманих за допомогою активних методів, можна розбити на два завдання: надроздільна реконструкція карти глибини (depth map super-resolution) та заповнення карти глибини на базі даних датчиків ToF і датчиків структурованого світла. Надроздільна реконструкція карти глибини фокусується на збільшенні просторової роздільної здатності карти глибини, а заповнення карти глибини фокусується на заповненні відсутніх значень глибини. Ці дві задачі схожі і можуть бути зведені до єдиного формулювання [83].

Методи удосконалення карт глибини, які не використовують зовнішні набори даних та методи, що використовують властивість співпадіння ребер на карті глибини з ребрами на відповідному кольоровому зображенні, можна визначити як методи на основі фільтрів та методи на основі оптимізації.

Методи на основі фільтрів працюють швидко, але їх можливість згладжування обмежена через локальний розв'язок. Однією з перших робіт, які присвячені методам на основі фільтрів є фреймворк спільної двосторонньої дискретизації (JBU) [84]. Згідно даної роботи, краї карти глибини низької роздільної здатності можуть бути уточнені відповідно до країв зареєстрованого кольорового зображення високої роздільної здатності за допомогою методів двосторонньої фільтрації.

у роботі [85] був запропонований варіант вищезазначеного фреймворку. Він обчислює ваги для усереднення на основі геодезичного простору, який є спільним простором кольору і відстані, замість того, щоб розділяти простір кольору і простір відстані. В роботі [86] автори уточнюють карти глибини ітеративно за допомогою об'єму витрат, який складається з набору кандидатів на значення глибини. Кожен елемент в об'ємі витрат обчислюється за допомогою JBU з певним кандидатом на значення глибини. Автори роботи [87] пропонують керовану фільтрацію зображень для покращення глибини. Запропонований метод моделює лінійну залежність між

вихідним і направляючим зображеннями, яка ґрунтується на припущенні, що вихід має ребро тільки тоді, коли вхід має ребро.

Автори роботи [88] пропонують метод зваженої фільтрації моди (WMF) на основі спільної гистограми кандидатів на значення глибини. Метод змушує результат задовольняти вимогі мінімізації за нормою  $L_1$ , яка є більш стійкою до викидів, ніж мінімізація за нормою  $L_2$ . Автори роботи [89] пропонують спільну тристоронню фільтрацію (JTF) для покращення глибини. Різниця між JBU та JTF полягає в тому, що цей фільтр інтегрує інформацію про локальний градієнт карти глибини низької роздільності. Автори роботи [90] пропонують підхід до надроздільної реконструкції карти глибини, який використовує процедуру фільтрації типу «лушпиння цибулі» для використання локальної інформації про градієнт карти глибини.

Порівняно з методами на основі фільтрів, методи на основі оптимізації є більш стійкими до шуму. Автори роботи [91] моделюють процес надроздільної реконструкції карти глибини як розв'язання задачі оптимізації з декількома мітками за допомогою Марківського випадкового поля. В роботі [92] запропоновано розширення попередньої роботи шляхом введення члена даних, який може краще відповідати характеристикам карт глибини. В роботі [93] автори оновлюють традиційне просторове Марківське випадкове поле до динамічного поля, яке розглядає як просторову, так і часову інформацію як складові функції енергії для підвищення точності та надійності надроздільної реконструкції карти глибини для динамічних сцен.

В роботі [94] автори пропонують нелокальний член для регуляризації карт глибини і поєднують його з ваговою схемою, яка включає інформацію про край, градієнт і сегментацію, витягнуту з кольорових зображень високої роздільності. Автори роботи [95] моделюють член регуляризації як загальне узагальнене варіаційне обмеження гладкості другого порядку і направляють процес надроздільної реконструкції карти глибини за допомогою анізотропного тензора дифузії, який обчислюється із зареєстрованого кольорового зображення високої роздільності. Автори роботи [96] досягають покращення глибини за допомогою

керованої кольором моделі авторегресії. Предиктор авторегресії для кожного пікселя будується відповідно до локальної кореляції на початковій карті глибини та нелокальної схожості на зареєстрованому високоякісному кольоровому зображенні. В роботі [97] автори розробляють базу термів регуляризації на основі надійного M-оцінювача, щоб неявно впоратися з неузгодженістю між картою глибини та зареєстрованим кольоровим зображенням.

### **3.1 Перешкоди в процесі керованого удосконалення карт глибини**

В основному, якість країв карти глибини низької роздільної здатності можна покращити, спираючись на вказівки, надані відповідними краями кольорового зображення. Однак припущення, що краї карти глибини та краї кольорового зображення у відповідних місцях узгоджуються, не завжди відповідає дійсності. Неправильне орієнтування на відповідне кольорове зображення призведе до появи артефактів копіювання текстури та розмиття меж глибини на покращеній карті глибини. У разі, якщо гладкі області отриманої карти глибини будуть відповідати областям кольорового зображення з багатою текстурою, це призводить до появи артефактів копіювання текстури. З іншого боку, розмиті краї карти глибини зазвичай спостерігаються у випадку, коли гладкі області кольорового зображення відповідають областям карти глибини з краями.

Існує декілька робіт, які намагаються знайти баланс між внеском оригінальної карти глибини та супутнього кольорового зображення для вирішення цих проблем, але такі методи мають два недоліки [90, 94, 96]:

1) вони не оцінюють в явному вигляді неузгодженість країв між кольоровим зображенням і відповідною картою глибини. Тому вони не можуть адаптивно керувати процесом зіставлення кольорового зображення при удосконаленні карти глибини;

2) схожість країв в енергетичній функції Марковського випадкового поля обчислюється незалежно та лише на основі різниці кольору та глибини між пікселем та його сусідніми пікселями на кольоровому зображенні та грубо

інтерпольованій карті глибини. Така схема обчислень ігнорує локальну структуру карти глибини.

В рамках даного дисертаційного дослідження пропонуються методи для вирішення вищезазначених недоліків. Зокрема, запропоноване жорстке та м'яке вимірювання неузгодженості країв має на меті вирішити перший недолік, а запропоноване використання мінімального покриваючого лісу має на меті вирішити другий недолік.

### **3.2 Керований метод покращення карти глибини з жорсткою оцінкою неузгодженості країв**

Щоб оцінити неузгодженість країв між картою глибини з низькою роздільною здатністю і кольоровим зображенням з високою роздільною здатністю, слід визначити кілька моментів:

– для того, щоб виміряти неузгодженість між картою глибини і кольоровою картою, роздільна здатність цих двох карт повинна бути однаковою. У цьому випадку, перед виявленням країв, карта глибини з нижчою роздільною здатністю або з отворами грубо інтерполюється до тієї ж роздільної здатності, що і кольорове зображення за допомогою методів сітчастої або розсіяної інтерполяції;

– оскільки кольорове зображення і відповідна карта глибини мають структурну подібність, яка чітко спостерігається на відповідних бінарних картах країв, запропонований метод вимірює неузгодженість між бінарними картами країв, згенерованими з кольорового зображення і відповідної карти глибини відповідно.

Для виявлення країв грубо інтерпольованої карти глибини та відповідного кольорового зображення використовується оператор Канні [98]. Через низьку роздільну здатність або шум, краї глибини можуть зміщуватися від своїх справжніх положень на карті глибини.

Таким чином, порівняно із зареєстрованими кольоровими краями, неузгоджені краї карти глибини можна розділити на два типи: один з них

погіршується грубою інтерполяцією, але може бути уточнений за допомогою наведення на високоякісні кольорові краї, інший – це неузгодженість країв.

Прямим рішенням є класифікація цих двох випадків для всіх крайових пікселів за допомогою жорсткого рішення, яке однозначно визначає реальний випадок.

Виходячи з того, що кількість помилок має бути меншою за можливості кодування з корекцією помилок, припускається, що край карти глибини погіршується грубою інтерполяцією, коли зміщення між ним і найближчим краєм кольорового зображення є меншим за порогове значення. В іншому випадку це трактується як справжня неузгодженість країв.

### 3.2.1 Визначення функції енергії Марківського випадкового поля

Відповідно до теореми Хаммерслі-Кліффорда [99], запропонований метод визначає цільову функцію для надроздільної реконструкції карти глибини наступним чином:

$$M' = \arg \min_{m_k \in M} \sum_{v_k \in V} \xi_c^k H_t(m_k, v_k) + \xi \sum_k \sum_{l \in N_k} \xi_f^{kl} H_r(m_k, m_l), \quad (3.1)$$

$$H_t(m_k, v_k) = |m_k - v_k|, \quad (3.2)$$

$$H_r(m_k, m_l) = |m_k - m_l|. \quad (3.3)$$

де  $M'$  – карта глибини,  $V$  – спостережені значення глибини,  $k, l$  – розташування пікселів на покращеній карті глибини,  $v_k$  – спостережуване значення глибини пікселя  $k$ ,  $N_k$  – множина 4-зв'язних сусідніх пікселів  $k$ ,  $H_t$  – член даних, який вказує на сумісність покращеної глибини зі спостережуваними значеннями,  $H_r$  – член регуляризації, який приводить до кусково-гладкого розв'язку і штрафує за різні значення глибини для сусідніх пікселів,  $\xi$  використовується для балансування члена даних і члена регуляризації.  $\xi_c^k$  представляє достовірність

спостережуваного значення глибини  $v_k$ ,  $\xi_f^{kl}$  позначає анізотропну спорідненість  $k, l$ , яка вбудовується в запропоноване вимірювання неузгодженості країв.

**Виявлення викидів.** Спостережувані значення глибин пікселів, розташованих на краях карти глибини низької роздільної здатності, є ненадійними, оскільки вони розмиті змішуванням значень глибин двох різних шарів глибин. Тому вони не повинні брати участь у побудові термів даних. У запропонованому методі для виявлення країв карти глибин низької роздільної здатності використовується простий оператор `sanny`.  $\xi_c^k$  – двійкова величина, яка дорівнює 0, якщо  $k$  є пікселем краю, інакше вона дорівнює 1.

**Обчислення анізотропної спорідненості.** Для кожного пікселя  $l$  по сусідству з пікселем  $k$  анізотропна спорідненість  $\xi_f^{kl}$  пари пікселів  $k, l$  визначається  $\xi_f^k$  і  $\xi_f^l$ , які пояснюються наступним чином:

– якщо на кольоровому зображенні та грубо інтерпольованій карті глибини піксель  $k$  постійно знаходиться на краю або на гладкій ділянці, то вважається, що розподіл кольорового зображення узгоджується з розподілом карти глибини поблизу пікселя  $k$ , а  $\xi_f^k$  може бути сформована як вагова функція, введена в JBU [84] для визначення ефекту уточнення карти глибини за допомогою кольорового зображення, який далі називається «Уточнення за кольоровим зображенням»;

– якщо  $k$  знаходиться на кольоровому зображенні на ребрі, але не на грубо інтерпольованій карті глибини, то на такій карті глибини визначається вікно пошуку. Параметру  $\xi_f^k$  може бути присвоєно значення «Уточнення за кольоровим зображенням», коли у вікні пошуку існують ребра, інакше параметр класифікується як ситуація істинної неузгодженості ребер. Оскільки  $k$  знаходиться в гладкій області на карті глибини, параметру  $\xi_f^k$  має бути присвоєно велике значення, щоб пригнітити різні призначення міток для сусідніх пікселів, які для стислості будуть називатися «Гладка область»;

– якщо  $k$  на грубо інтерпольованій карті глибини знаходиться на краю, але не на кольоровому зображенні, визначається вікно пошуку на кольоровому зображенні. Параметру  $\xi_c^k$  може бути присвоєно значення «Уточнення за

кольоровим зображенням», якщо у вікні пошуку існують краї, інакше це класифікується як ситуація істинної неузгодженості країв. Оскільки  $k$  знаходиться поблизу ребер глибини, параметру  $\xi_c^k$  має бути присвоєно мале значення, щоб підвищити різні призначення міток для сусідніх вузлів, які для стислості називаються «Поблизу країв».

Згідно з наведеними вище умовами,  $\xi_f^k$  та  $\xi_f^l$  обчислюються таким чином:

$$\xi_f^{\{k;l\}} = \begin{cases} e^{-\frac{\Delta L_{kl}^2}{\sigma^2}}, & \text{"Уточнення за КЗ"} \\ e^{-\frac{\Delta L_{sm}^2}{\sigma^2}}, & \text{"Поблизу країв"} \\ e^{-\frac{\Delta L_{lg}^2}{\sigma^2}}, & \text{"Гладка область"} \end{cases} \quad (3.4)$$

де  $\Delta L_{kl}$  – різниця яскравості пари пікселів  $k, l$ ;  $\Delta L_{sm} = 1$ , а  $\Delta L_{lg} = 254$ . Через симетричне співвідношення  $k, l$  в парі пікселів, параметр  $\xi_f^{kl}$  визначається за допомогою  $\xi_f^k$  та  $\xi_f^l$ .

Фактично, якщо  $\xi_f^k$  та  $\xi_f^l$  приймають однакове рішення,  $\xi_f^{kl}$  визначається без двозначності. Крім того, згідно з визначенням, помічено, що якщо  $\xi_f^k$  класифікується як «Гладка область», то  $\xi_f^l$  не може бути класифікована як «Поблизу країв», і навпаки.

Для решти ситуацій, якщо  $\xi_f^k$  класифікується як «Уточнення за кольоровим зображенням», а  $\xi_f^l$  – як інша категорія, показано, що  $l$  знаходиться біля краю або кольорового зображення, або карти глибини, отже,  $\xi_f^{kl}$  визначається відповідно до  $\xi_f^l$ .

Всі ситуації наведено в таблиці 3.1, де набори значень  $\xi_f^k$  та  $\xi_f^l$  наведено у першому стовпчику та рядку відповідно, а кожна комірка містить значення  $\xi_f^{kl}$  при заданому значенні  $\xi_f^k$  та  $\xi_f^l$ .

Таблиця 3.1 – Визначення параметру  $\xi_f^{kl}$  на базі значень  $\xi_f^k$  та  $\xi_f^l$ 

Значення $\xi_f^k$	Значення $\xi_f^l$		
	«Уточнення за КЗ»	«Поблизу країв»	«Гладка область»
«Уточнення за КЗ»	«Уточнення за КЗ»	«Поблизу країв»	«Гладка область»
«Поблизу країв»	«Поблизу країв»	«Поблизу країв»	---
«Гладка область»	«Гладка область»	---	«Гладка область»

### 3.2.2 Експериментальне дослідження запропонованого методу

Для оцінки ефективності запропонованого методу використано набір стереоданих Middlebury [73], який містить пари RGB-D зображень. Карти глибини високої роздільної здатності продискретизовано за чотирма факторами (2×, 4×, 8× та 16×) для формування вхідної карти глибини низької роздільної здатності. Запропонований метод порівнюється з бікубічною інтерполяцією, оригінальним методом на основі Марківського випадкового поля [100], спільною двосторонньою дискретизацією [84], надроздільною реконструкцією карти глибини з використанням геодезичного простору [101], методом керованої фільтрації зображень [87], методом глобальної узагальненої дисперсії [95] та методом рухомих найменших квадратів [102].

Об'єктивні результати, оцінені за допомогою середнього абсолютного відхилення, наведені в таблиці 3.2.

Загалом, запропонований метод може забезпечити задовільні результати для невеликого коефіцієнта збільшення вибірки. Однак, при більшому коефіцієнті дискретизації він не є надійним. Зі збільшенням коефіцієнта дискретизації якість карт країв значно погіршується.

Таким чином, показано недолік вимірювання неузгодженості країв з жорсткою дискретністю. Порівняно із запропонованим методом, метод рухомих найменших квадратів та метод глобальної узагальненої дисперсії показують кращі результати у випадку великих коефіцієнтів дискретизації (8× та 16×).

Таблиця 3.2 – Результати роботи методів покращення карти глибини

Метод	Набір даних															
	Book				Moebius				Dolls				Reindeer			
	2x	4x	8x	16x	2x	4x	8x	16x	2x	4x	8x	16x	2x	4x	8x	16x
Guided	0,23	0,36	0,59	1,15	0,24	0,39	0,61	1,17	0,29	0,36	0,57	1,15	0,43	0,56	0,89	1,82
JBU	0,18	0,37	0,75	1,58	0,19	0,38	0,77	1,48	0,22	0,40	0,76	1,48	0,28	0,52	1,02	1,90
TGV	0,20	0,28	0,43	0,84	0,21	0,30	0,51	0,89	0,23	0,39	0,72	2,21	0,33	0,51	1,05	3,07
Bicubic	0,16	0,30	0,61	1,16	0,14	0,32	0,61	1,15	0,22	0,40	0,68	1,20	0,32	0,57	1,01	1,89
MLS	0,17	0,28	0,47	1,18	0,17	0,36	0,51	0,94	0,25	0,40	0,62	0,99	0,34	0,65	0,77	1,45
Власний	0,14	0,26	0,49	0,93	0,17	0,28	0,64	1,19	0,19	0,38	0,74	1,45	0,23	0,42	0,77	1,52

### 3.3 Керований метод удосконалення карти глибини з м'якою оцінкою неузгодженості країв

Через недоліки вимірювання неузгодженості країв з жорстким рішенням, у цьому підрозділі пропонується метод вимірювання неузгодженості країв з м'яким рішенням, який точніше вимірює ступінь неузгодженості, що виникає між краями карти глибини та відповідними краями кольоровому зображенні в числовому вигляді. Таке вимірювання неузгодженості країв вбудовано в Марківське випадкове поле, що дає змогу зменшити артефакти копіювання текстури та зберегти краї карти глибини, порівняно з аналогом з жорстким рішенням.

#### 3.3.1 Модифікована функція енергії Марківського випадкового поля

Оскільки багатоміткова оптимізація за допомогою зрізів на графах не може точно досягти глобального мінімуму, а значення глибини завжди записується в міліметрах як безперервна плаваюча величина, наприклад, у наборах даних ToF-Mark [95], у цьому підрозділі пропонується використання Марківського випадкового поля з безперервними змінними. Для спрощення оптимізації енергетична функція модифікується у вигляді наступного рівняння:

$$M' = \arg \min_{m_k \in M} \sum_{v_k \in V} \xi_c^k H_t(m_k, v_k) + \xi \sum_k \sum_{l \in N_k} \xi_f^{kl} H_r(m_k, m_l), \quad (3.5)$$

$$H_t(m_k, v_k) = (m_k - v_k)^2, \quad (3.6)$$

$$H_r(m_k, m_l) = (m_k - m_l)^2. \quad (3.7)$$

де  $\xi_f^{kl}$  обчислюється на основі запропонованого вимірювання неузгодженості країв з м'яким прийняттям рішень.

Основні відмінності між цим підходом та підходом, представленим в попередньому підрозділі, полягають у двох аспектах:

- 1) припускаючи наявність гаусівського шуму у необробленій карті глибини низької роздільності, член даних і член регуляризації визначаються у квадратичній функції замість абсолютного аналога для кращої продуктивності згладжування;
- 2) через недоліки вимірювання неузгодженості країв з жорстким прийняттям рішень, спорідненість члена регуляризації в цьому підході обчислюється за допомогою вимірювання неузгодженості країв з м'яким прийняттям рішень.

### 3.3.2 Вимірювання неузгодженості країв з м'яким прийняттям рішень

Згідно з роботою [103], вимірювання неузгодженості країв між кольоровим зображенням та картою глибини можна представити як двонаправлену оцінку якості карти країв. Однак, згідно вищезазначеної роботи, загальне вимірювання якості карти країв ґрунтується на зміщенні положення кожного пікселя краю відносно положення у еталонних даних.

В даному випадку, коли відбувається покращення карти глибини, пікселі країв на карті глибини і кольоровому зображенні, які мали б знаходитися в однаковому положенні, завжди мають зміщення один щодо одного. Причинами цього є певна попередня обробка, наприклад, груба інтерполяція або шуми в датчиках глибини. Таким чином, неможливо виміряти неузгодженість за

допомогою різниці між положеннями кожної пари співставлених пікселів країв, як це роблять існуючі методи вимірювання якості краю [104]. Натомість, запропонований метод вимірювання неузгодженості країв базується на подібності структури карт країв, яка враховує структуру, представлену локальними сусідніми регіонами, а також глобальну структуру всієї карти країв.

Для кожного пікселя на еталонній карті країв відбувається пошук найкращої відповідності на цільовій карті країв в околиці навколо відповідної позиції. Це означає, що якщо край кольорового зображення і край карти глибини узгоджуються, то зміщення пікселів країв, що співпадають, має бути обмежене в невеликому діапазоні. Крім того, величина та орієнтація зсувів усіх пікселів країв у сусідній області мають бути однаковими. Ці два обмеження розв'язуються в оптимізації Марківського випадкового поля за допомогою члена даних і члена регуляризації відповідно, як показано в рівнянні (3.8). Член даних містить інформацію про локальну структуру, а член регуляризації – інформацію про глобальну структуру. Таким чином, вимірювання неузгодженості країв є стійким до помилок у вихідній глибинній карті країв:

$$F^* = \arg \min_{f \in F} \sum_{k \in ref} S(k, k + f_k) + \mu \sum_{k \in ref} \sum_{l \in N_k} T(f_k, f_l), \quad (3.8)$$

де  $S$  і  $T$  – функції, визначені для членом даних і термів регуляризації в Марківському випадковому полі відповідно,  $\mu$  – коефіцієнт балансу між членом даних і членом регуляризації (у запропонованому методі він дорівнює 0,1),  $k$  – позиція пікселя краю на еталонній карті країв  $ref$ ,  $N_k$  – множина 8-зв'язних сусідніх пікселів  $k$ ,  $f_k$ , який є елементом  $L$ , позначає зміщення для  $l$ . Таким чином,  $k + f_k$  представляє позицію пікселя краю  $k$  на цільовій карті країв.

Оскільки підпікселі, створені віртуально в процесі інтерполяції, можуть бути нестабільними, кожен піксель краю на еталонній карті країв зіставляється з існуючим пікселем краю, виявленим на цільовій карті країв. Іншими словами,

запропонований метод призначає  $F$  з точністю до цілого пікселя, і це є дискретною задачею оптимізації Марківського випадкового поля.

У всіх експериментах розмір вікна пошуку визначає діапазон  $f_k$ , який становить  $5 \times 5$  для  $2 \times$  надроздільної реконструкції,  $7 \times 7$  для  $4 \times$  надроздільної реконструкції,  $9 \times 9$  для  $8 \times$  надроздільної реконструкції,  $11 \times 11$  для  $16 \times$  надроздільної реконструкції і  $7 \times 7$  для заповнення карти глибини без надроздільної реконструкції.

Терм даних  $S(k, l)$  – це вартість зіставлення референсного пікселя краю  $k$  з цільовим пікселем краю  $l$ . При заданому  $k$ , якщо відповідний цільовий піксель  $l$  на цільовій карті країв не є пікселем краю, це вважається певною неузгодженістю при поточному  $f_k$ . У цьому випадку  $S(k, l)$  присвоюється максимальне значення неузгодженості (тобто 1 у запропонованому методі). В іншому випадку ця неузгодженість вимірюється на двох ділянках, де піксель краю  $k$  і піксель краю  $l$  є центральними позиціями відповідно.

У запропонованому методі розмір ділянки становить  $3 \times 3$ . Це вимірювання відбувається за допомогою методу мінімально-зваженого двостороннього збігу [105], який є більш надійним, ніж метод середнього абсолютного відхилення. Насправді, мінімально зважений двосторонній збіг базується на схожості структури. Однак метод середнього абсолютного відхилення враховує лише різницю кожної пари пікселів незалежно. У зваженому двосторонньому графі кожне ребро графа має відповідне значення. Мінімальне зважене двостороннє зіставлення полягає в тому, щоб знайти найкраще зіставлення, де сума значень ребер графа, що з'єднують вершини, які зіставляються, є мінімальною.

У запропонованому методі якість двостороннього зіставлення вимірюється за різницею між розташуванням пікселів країв, що зіставляються, та сумою пікселів країв у двох ділянках. На рисунку 3.1 наведено ілюстрацію переваги зваженого двостороннього зіставлення порівняно із методом середнього абсолютного відхилення. На трьох ділянках, показаних на рисунку, білі та чорні пікселі представляють крайові та некрайові пікселі відповідно. При застосуванні

середнього абсолютного відхилення і (b), і (c) (тобто цільові ділянки) мають однакову вартість зіставлення з еталонною ділянкою (a). Однак можна помітити, що з точки зору локальної структури цільова ділянка (b) ближча до а). Таку дрібнозернисту схожість можна успішно вловити за допомогою двостороннього зіставлення графів, що використовується в запропонованому методі.

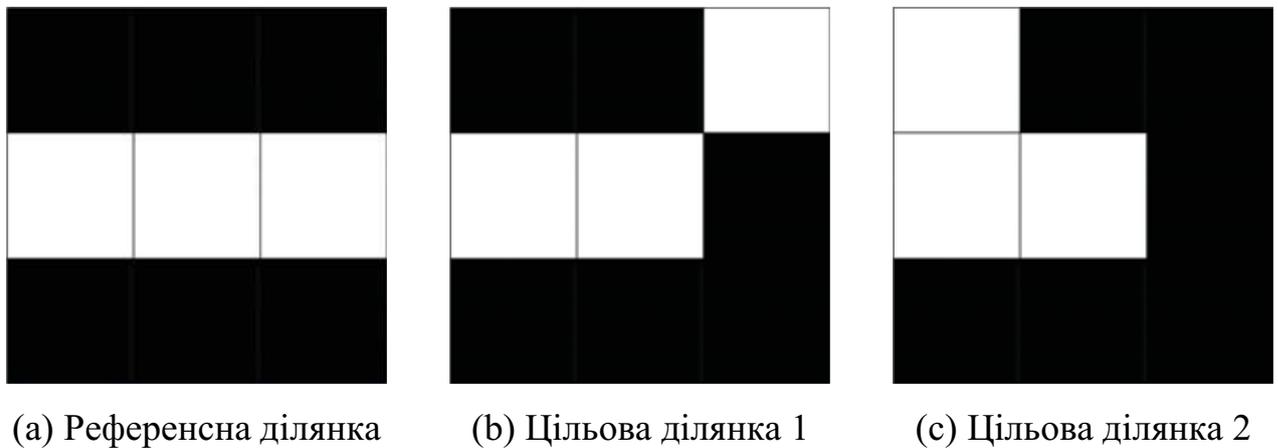


Рисунок 3.1 – Ілюстрація переваги методу зваженого двостороннього зіставлення

Виходячи з аналізу, наведеного вище, член даних  $S(k, l)$  виражається у вигляді наступного рівняння:

$$S(k, l) = \begin{cases} 1(\text{безумовна невідповідність}), & \text{якщо } l \notin \text{пікселі краю} \\ MT(V_k, V_l, E, VEC) & \text{, в іншому випадку} \end{cases} \quad (3.9)$$

де  $MT$  означає мінімально зважене двостороннє зіставлення. Дводольний граф  $G(V_k, V_l, E, VEC)$  визначається наступним чином;  $V_k$  та  $V_l$  – це вершини,  $E$  – ребра графа між вершинами, а  $VEC$  – вектор, який присвоює вагу кожному ребру графа в  $E$ . Зокрема,  $V_k = \{ek_1, ek_2, \dots, ek_m\}$  і  $V_l = \{el_1, el_2, \dots, el_n\}$  представляють множини пікселів країв у двох ділянках (за винятком  $k$  і  $l$ , які є центральними пікселями країв цих двох ділянок),  $m$  і  $n$  – кількість пікселів країв усередині цих двох множин відповідно.

Таким чином, вимірювання неузгодженості між  $k$  та  $l$  розглядається як проблема узгодження двох наборів даних  $V_k$  та  $V_l$ . Крім того, припускається, що

розташування крайнього пікселя та його істинного крайнього пікселя є близькими один до одного. Це припущення відповідає подібності локальної структурної інформації. Тому кожен елемент  $VEC$  визначається як  $\varepsilon(ek_i, el_j)$ , яка є монотонною функцією, що повертає позитивний штраф за локальну структурну відповідність:

$$\varepsilon(ek_i, el_j) = f(|ek_i^x + el_j^x| + |ek_i^y - el_j^y|) \quad (3.10)$$

де  $f(0) = 0$ ,  $f(1) = 1$ ,  $f(2) = 1,6$  і  $f(x) = 2$ , якщо  $x > 2$ ,  $ek_i^x$  і  $ek_j^x$  – координати пікселя краю  $ek_i$  відповідно.

Мінімально-зважене двостороннє зіставлення використовується для забезпечення відповідності один до одного між наведеними вище наборами даних пікселів країв. Це означає, що будь-який піксель краю у  $V_k/V_l$  збігається щонайбільше з одним крайнім пікселем у  $V_l/V_k$  з  $|m - n|$  неспівпадаючими пікселями.

На рисунку 3.2 наведено ілюстрацію мінімально-зваженого двостороннього зіставлення з позначеними неузгодженими пікселями. Насправді, кількість пікселів, що не співпадають, також відображає структурні відмінності між наборами пікселів країв  $V_k$  і  $V_l$ .

Крім того, щоб ефективно зменшити вплив помилок виявлення країв на зашумлених картах глибини, слід враховувати різницю, відображену цими неспівпадаючими пікселями.

В процесі дослідження роботи методу було зроблено спостереження, що коли кількість пікселів країв на обох ділянках сильно відрізняється, запропонований метод вважає, що цей край може бути спричинений шумом або що таке зіставлення не є надійним.

Таким чином, він повинен додавати великі витрати на таке узгодження. Щоб врахувати ці проблеми, член вимірювання неузгодженості країв  $MT(V_k, V_l, E, VEC)$  у рівнянні (3.5) скориговано і визначено у вигляді наступного рівняння:

$$MT(V_k, V_l, E, VEC) = \left( \sum_{(uk_h, ul_h) \in V'_{kl}} \frac{\varepsilon(uk_h, ul_h)}{2} + |m - n| \right) / 8, \quad (3.11)$$

де  $V'_{kl} = \{(uk_1, ul_1), (uk_2, ul_2), \dots, (uk_r, ul_r)\}$  – множина пар пікселів країв, обраних методом мінімально зваженого двостороннього зіставлення,  $\varepsilon(uk_h, ul_h)$  – вага ребра, що з'єднує пікселі країв  $uk_h$  та  $ul_h$ , з  $h = \{1, 2, \dots, r\}$ . Отже,  $\sum_{(uk_h, ul_h) \in V'_{kl}} \varepsilon(uk_h, ul_h)$  – це вартість зіставлення мінімально зваженого двостороннього зіставлення, згаданого вище. Для того, щоб обмежити терм даних  $S(k, l)$  в діапазоні  $[0, 1]$ , до рівняння (3.7) застосовується звичайна нормалізація

$T(f_k, f_l)$  – це член нормалізації в рівнянні (3.4), який дає штраф, коли сусідні пікселі країв мають різні зміщення:

$$T(f_k, f_l) = \begin{cases} 0, & \text{якщо } f_k = f_l \\ 1, & \text{в іншому випадку} \end{cases} \quad (3.12)$$

де  $f_k$  – вектор зміщення для пікселя краю  $k$ .

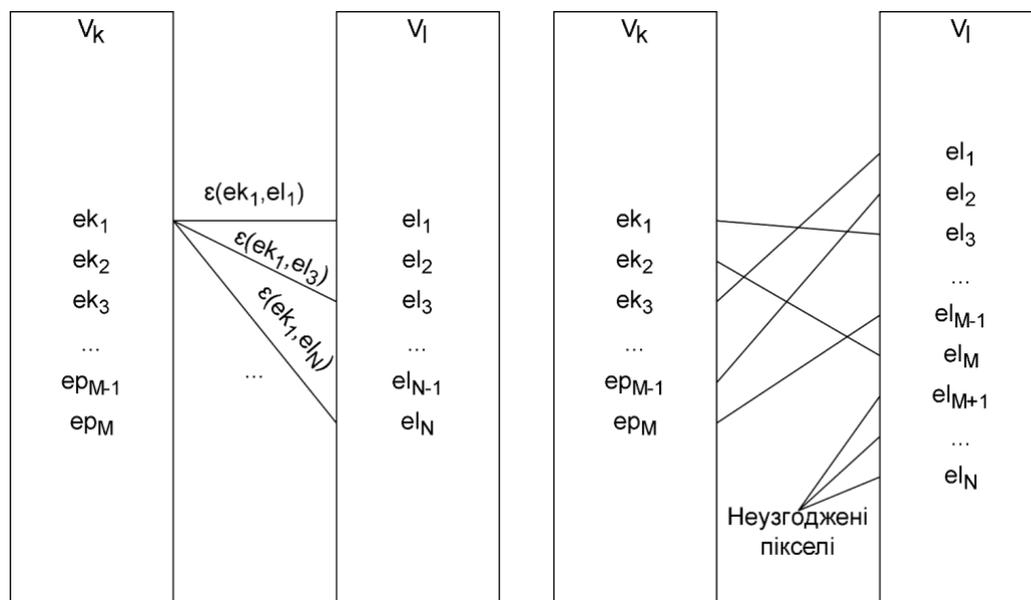


Рисунок 3.2 – Ілюстрація задачі мінімального зваженого двостороннього зіставлення з конфігурацією зваженого дводольного графа та (b) результатом роботи методу з неузгодженими пікселями

Після того, як член даних і член нормалізації в рівнянні (3.4) визначені, для розв'язання задачі багатокритеріальної оптимізації Марківського випадкового поля застосовується метод розрізу графа. Тоді вимірювання неузгодженості для пікселя краю  $k$ , представленого  $S(k, l)$ , можна обчислити за оптимізованим зміщенням  $f'_k \in F'$ . У запропонованому методі карта неузгодженостей  $S_{ref}$  – це множина  $S(k, k + f'_k)$ , що представляє вимірювання неузгодженостей для всіх пікселів країв на еталонній карті країв. Якщо для певного пікселя краю не знайдено збігів, то зміщення  $f$  цього крайового пікселя не має сенсу, а його значення неузгодженості набуває максимального значення 1.

Неузгодженість країв вимірюється на основі еталонної карти країв у порівнянні з цільовою картою країв. Таким чином, результати вимірювання будуть відрізнятися при заміні цих двох карт країв. У цій роботі дві карти країв – це карта країв кольорового зображення і відповідна карта країв карти глибини. Коли карта країв кольорового зображення розглядається як еталонна карта країв, можна помітити, що виявлені найбільш неузгоджені позиції відображають області, де відбувається копіювання текстури.

З іншого боку, коли карта країв карти глибини розглядається як еталонна карта країв, можна помітити, що найбільш неузгоджені позиції відображають області розмиття країв карти глибини. На рисунку 3.3 показано вимірювання двонаправленої неузгодженості для набору Art з набору даних Middlebury, яка виражається у помилкових кольорових зображеннях. На рис. 3.3, в і рисунку 3.3, г колір уздовж пікселів країв відображає величину неузгодженості країв між еталонною картою країв і цільовою картою країв.

Після двонаправленої оцінки формуються дві карти неузгодженості  $S_{clr}$  (карта країв кольорового зображення, яка вважається еталонною),  $S_{dm}$  (карта країв глибинного зображення вважається цільовою), а також два набори зміщень  $F_{clr}$ ,  $F_{dm}$  для пари зображень. Перед вбудовуванням значень вимірювань неузгодженостей у запропонований метод на основі Марківського випадкового поля, ці дві карти неузгодженостей мають бути об'єднані одна з одною.



(а)



(б)



(в)



(г)

Рисунок 3.3 – Вимірювання двонаправленої неузгодженості для набору даних Міддлбері «Art»: (а) карта країв кольорового зображення, (б) карта країв глибини, (в) вимірювання неузгодженості країв у випадку, коли карта країв кольорового зображення є еталонною картою країв, (г) вимір неузгодженості країв, у випадку, коли карта країв глибини є еталонною картою країв

Як згадувалося раніше, положення пікселів країв на грубо інтерпольованій карті глибини є ненадійними. Навпаки, положення пікселів країв на кольоровій карті країв є більш точним завдяки високій якості кольорового зображення. Розв'язуючи задачу оптимізації Марківського випадкового поля у рівнянні (3.4) з

використанням карти країв глибинного зображення як еталонної карти країв, зміщення між кожним пікселем краю  $k$  глибинного зображення і відповідним йому пікселем краю  $k$  кольорового зображення становить  $S_{dm}(k)$ . Отже, істинне розташування пікселя краю зображення карти глибини  $k$ , що спостерігається, буде ближчим до  $k + F_{dm}(k)$  коли  $S_{dm}(k) \neq 1$ .

У випадку однозначної неузгодженості  $S_{dm}(k) = 1$ , положення пікселя краю  $k$  залишається незмінним, оскільки на карті країв кольорового зображення немає відповідного пікселя краю.

Більше того, через невизначеність зміщень, пікселю краю кольорового зображення  $k'$  може відповідати більше одного пікселя краю карти глибини  $k$ . У цій ситуації значення неузгодженості оновлюється для найкращого положення пікселя краю з найнижчою вартістю відображення. У запропонованому методі скоригована глибина  $S'_{dm}$  виражається як:

$$S'_{dm}(k') = \min_{k \in \{k | k' = F_{dm}(k)\}} S_{dm}(p), \quad \text{якщо } S_{dm}(k) \neq 1 \quad (3.13)$$

$$S'_{dm}(k) = S_{dm}(k), \quad \text{в іншому випадку} \quad (3.14)$$

Після вирівнювання двох карт неузгодженості  $S'_{dm}$  і  $S_{clr}$ , визначається довірча карта  $\lambda$ , яка враховує вимірювання неузгодженості у двонаправленому обчисленні.

Вона описує остаточний стан неузгодженості між картою границь кольорового зображення і картою границь глибинного зображення. Довірча карта обчислюється за формулою:

$$\lambda = \max(S'_{dm}, S_{clr}). \quad (3.15)$$

Далі треба розглянути процес застосування отриманих значень в Марківському випадковому полі для уточнення ступеня орієнтування на кольорове зображення в процесі керованого удосконалення карти глибини.

### 3.3.3 Включення вимірювань неузгодженості країв у Марківське випадкове поле

Взагалі кажучи, інформація для керованого удосконалення карти глибини може бути отримана з двох джерел. Перше – з отриманого кольорового зображення, а друге – з самої вихідної карти глибини. На основі карти впевненості  $\lambda$ , яка обчислюється в рівнянні (3.11), запропонований метод поєднує ці два види інформації для створення нових вказівок для обчислення спорідненості  $\xi_f^{kl}$ :

$$\xi_f^{kl} = e^{-\frac{(|\Delta_c^{kl}|(1-\lambda_{kl})+|\Delta_d^{kl}|\lambda_{kl})^2}{2\delta^2}} \quad (3.16)$$

де  $\Delta_c^{kl}$  і  $\Delta_d^{kl}$  представляють різницю кольорів і різницю глибин між позицією  $k$  і сусіднім пікселем  $l$  на кольоровому зображенні і грубо інтерпольованій карті глибини відповідно,  $\delta$  задає ширину смуги експоненціальної функції. Крім того, як згадувалося вище, довірчі значення доступні лише для крайніх пікселів. Для обчислення робастного довірчого значення для кожної пари пікселів  $k, l$ , застосовується операція максимуму ( $\lambda(k), \lambda(l)$ ), яка вибирає більш неузгоджене між довірчими значеннями  $k$  і  $l$ .

Це може зменшити артефакти копіювання текстури та краще зберегти краї карти глибини завдяки краям шириною в один піксель, виявленим оператором Кенні. Значення  $\lambda_{kl}$  виражається як  $\lambda_{kl} = \max(\lambda(k), \lambda(l))$ . Зокрема, коли сусідні пари пікселів  $k, l$  розташовані на краях як на кольоровому зображенні, так і на карті глибини,  $\lambda_{kl}$  буде мати значення близьке до 0, а  $\Delta_c^{kl}$  відіграє більш важливу роль в обчисленні подібності  $\xi_f^{kl}$ .

У такій ситуації, орієнтування на кольорове зображення допомагає відновити гострі краї на покращеній карті глибини. На противагу цьому, коли сусідня пара пікселів  $k, l$  розташовані на краю або кольорового зображення, або карти глибини,

але не на обох,  $\lambda_{kl}$  буде мати значення ближче до 1, а основне керування буде приходити від  $\Delta_d^{kl}$ .

У цих двох ситуаціях удосконалення глибини відбувається за допомогою методу покращення однієї карти глибини. Дійсно, можна застосувати певний метод покращення однієї карти глибини, щоб отримати більш точну карту глибин замість простої інтерпольованої карти глибин. Однак, покращення не є значним, коли коефіцієнт дискретизації малий. З іншого боку, використовуючи методи покращення однієї карти глибини, важко отримати точні межі глибин для великих коефіцієнтів дискретизації. Тому, запропонований метод використовує грубу інтерпольовану карту глибин як джерело орієнтування для всіх випадків. Перевага використання  $\Delta_d^{kl}$  є подвійною. З одного боку, це зменшує артефакти копіювання текстури. З іншого боку, орієнтація на інтерпольовану карту глибини є більш обґрунтованим, ніж помилкова орієнтація на кольорове зображення.

Сценарій, описаний вище, стосується областей навколо пікселів країв. Для пікселів, розташованих на гладких ділянках, де немає пікселів країв ні на кольоровому зображенні, ні на грубо інтерпольованій карті глибини, рівняння (3.12) не може бути застосованим для цього випадку, оскільки неможливо обчислити неузгодженість країв в локальній області, де пікселів країв немає взагалі. У запропонованому методі, для такого випадку використовується наступне рівняння:

$$\xi_f^{kl} = e^{-\frac{(|\Delta_d^{kl}|)^2}{2\delta^2}}. \quad (3.17)$$

В цьому випадку інформація для покращення глибини береться лише з грубо інтерпольованої карти глибини, щоб зменшити артефакти копіювання текстури. Виходячи з наведеного вище аналізу, запропонований метод може ефективно зберігати краї карти глибини і зменшувати артефакти копіювання текстури, адаптивно керуючи ступенем орієнтації на кольорове зображення.

Крім того, в областях поблизу країв глибини  $\delta$  має бути малим для збереження країв глибини. І навпаки,  $\delta$  має бути великим, щоб подавляти шум на гладких ділянках. Запропонований метод призначає різні значення  $\delta$  для гладких і негладких областей відповідно, які визначаються картою країв карти глибини. Зокрема якщо на карті країв глибини, якщо в локальних вікнах з центром в точці  $k$  і сусідньому пікселі  $l$  немає пікселя краю, то пара пікселів  $k, l$  знаходиться на гладкій ділянці. В іншому випадку така пара пікселів знаходиться на негладкій ділянці. У запропонованому методі,  $\delta$  дорівнює 2 і 4 для негладких областей (рівняння (3.12)) і гладких областей (рівняння (3.13)) відповідно.

### 3.3.4 Дослідження складності методу

На етапі вимірювання неузгодженості країв, задача розрізу на графах з декількома мітками розв'язується декількома підзадачами розрізу на графах з двома мітками за допомогою методу  $\alpha$ -розкладання. Складність задачі розрізу на графах з двома мітками досягає  $O(MN^2|C|)$ , де  $M$  і  $N$  – кількість ребер і вершин графа (тобто кількість пікселів країв, виявлених на еталонній карті країв) у графі відповідно; а  $|C|$  – це вартість мінімального розрізу, тобто найменша сумарна вага ребер, видалення яких від'єднає джерело від стоку. Таким чином, складність розрізу графа з декількома мітками досягає  $O(LMN^2|C|)$ , де  $L$  – кількість міток. Крім того, складність угорського алгоритму для мінімально-зваженого двостороннього зіставлення графів становить  $O(V^2E)$ , де  $V$  і  $E$  – кількість вершин (тобто, кількість країв ребер у двох ділянках) і ребер графа відповідно.

### 3.3.5 Експериментальне дослідження запропонованого методу

Експериментальне дослідження запропонованого методу проводилося на апаратній платформі з процесором Intel i5-14400F та 16 Гігабайтами оперативної пам'яті. Запропонований метод був реалізований на мові C, включаючи метод

розрізу графа. Середній час роботи методу склад 6,7 секунд при підвищенні частоти дискретизації з коефіцієнтом 16x карти глибини низької якості до роздільної здатності 1376x1088. Час виконання кожного кроку методу наведений в таблиці 3.3.

Експериментальне дослідження складається з трьох експериментів:

– перший експеримент полягає в оцінці ефективності запропонованого методу на наборі даних Middlebury, в яких синтетичні карти глибин деградували вручну різними способами. В рамках даного експерименту наведено порівняння продуктивності запропонованого методу з кількома існуючими методами;

– другий експеримент полягає в застосуванні запропонованого методу на реальних наборах даних ToF-Mark[95] та NYU [106] для отримання високоякісних карт глибини, щоб показати робастність запропонованого методу на реальних сценах;

– третій експеримент полягає в демонстрації продуктивності запропонованого методу для покращення карти глибини, яка полягає у вирішенні складної ситуації, коли відбувається комплексна деградація. Це стосується як низької роздільної здатності, так і значних прогалин в карті глибини.

Таблиця 3.3 – Середній час роботи запропонованого методу (16×)

Оцінка неузгодженості	Оптимізація МВП	Всього
18,14 сек.	14,33 сек.	32,47 сек.

Стосовно параметру  $\xi$  у рівнянні (3.3), то його можна теоретично проаналізувати у двох аспектах, які ґрунтуються на коефіцієнті дискретизації та наявності шуму на карті глибин низької роздільної здатності. З одного боку,  $\xi$  має зменшуватися зі збільшенням коефіцієнта дискретизації. Більший коефіцієнт дискретизації призведе до розрідженості набору пікселів, які мають спостережувані значення глибини, тому внесок члена даних у Марківське випадкове поле є незначним. Щоб збалансувати внесок члена даних і члена

регуляризації, необхідно зменшити  $\xi$ , таким чином, внесок члена даних буде відносно більшим, навіть якщо спостережувані дані про глибину є розрідженими. З іншого боку, збільшення  $\xi$  у випадку сильнішого шуму може забезпечити більшу стійкість моделі Марківського випадкового поля до шуму за рахунок посилення впливу члена регуляризації.

В процесі експериментального дослідження була виявлена закономірність, що коли карта глибини, отримана за допомогою інтерполяції за методом найближчого сусіда, не містить шуму, коефіцієнт дискретизації має менший вплив на  $\xi$ . Тобто, для різних коефіцієнтів дискретизації оптимальне  $\xi$  може мати близькі значення, доки шум на карті глибин не є значним.

За допомогою процесу перехресної перевірки, у запропонованому методі значення  $\xi$  встановлене на рівні 0,01 для всіх експериментів, в яких карти глибин низької роздільної здатності беруться з бенчмарку Middlebury без додавання шуму. У випадку додавання шуму, зв'язок між  $\xi$  і коефіцієнтом дискретизації визначається як:

$$\xi = \frac{const}{factor}, \text{ якщо } factor > 1 \quad (3.18)$$

де  $const$  – константа, яка дорівнює 3,2 у всіх експериментах з картами глибини низької роздільної здатності,  $factor$  – коефіцієнт дискретизації. Крім того, значення  $\xi$  дорівнювалося для всіх експериментів із побудови карти глибини.

### 3.3.5.1 Експеримент на наборі даних із синтетичними деградаціями

У даному експерименті було використано чотири дані з бенчмарку Middlebury: «Book», «Moebius», «Dolls» та «Reindeer».

В експерименті були розглянуті три типи деградації:

- 1) зменшення вибірки;
- 2) зменшення вибірки з додаванням шуму;

3) структурна помилка та випадкові пропуски.

**Деградація шляхом зменшення вибірки.** Запропонований метод порівнювався з шістьма методами-аналогами (Метод «Власний-Ж» означає варіант запропонованого методу з жорстким прийняттям рішень). В таблиці 3.4 наведені результати апсемплінгу за чотирма різними коефіцієнтами дискретизації. Помітно, що запропонований метод дає найнижче значення середнього абсолютного відхилення для більшості випадків. У випадку коефіцієнту дискретизації 16×, грубо збільшена роздільна здатність карти глибини вносить значні похибки, що впливає на якість карти країв глибини.

Таблиця 3.4 – Результати роботи методів на вибірці Middlebury, без шуму

Метод	Набір даних															
	Book				Moebius				Dolls				Reindeer			
	2x	4x	8x	16x												
Guided	0,23	0,36	0,59	1,15	0,24	0,39	0,61	1,17	0,29	0,36	0,57	1,15	0,43	0,56	0,89	1,82
JBU	0,18	0,37	0,75	1,58	0,19	0,38	0,77	1,48	0,22	0,40	0,76	1,48	0,28	0,52	1,02	1,90
TGV	0,20	0,28	0,43	0,84	0,21	0,30	0,51	0,89	0,23	0,39	0,72	2,21	0,33	0,51	1,05	3,07
Bicubic	0,16	0,30	0,61	1,16	0,14	0,32	0,61	1,15	0,22	0,40	0,68	1,20	0,32	0,57	1,01	1,89
MLS	0,17	0,28	0,47	1,18	0,17	0,36	0,51	0,94	0,25	0,40	0,62	0,99	0,34	0,65	0,77	<b>1,45</b>
Власний-Ж	0,14	0,26	0,49	0,93	0,17	0,28	0,64	1,19	0,19	0,38	0,74	1,45	0,23	0,42	0,77	1,52
Власний	<b>0,09</b>	<b>0,19</b>	<b>0,38</b>	<b>0,73</b>	<b>0,09</b>	<b>0,19</b>	<b>0,39</b>	<b>0,83</b>	<b>0,13</b>	<b>0,28</b>	<b>0,49</b>	<b>0,86</b>	<b>0,18</b>	<b>0,33</b>	<b>0,63</b>	1,46

Однак, продуктивність запропонованого методу в цьому випадку досягає найкращих показників у 3 з 4 випадків та субоптимального результату в четвертому випадку. Показано, що запропонований метод є робастним до якості карти країв глибини. На рисунку 3.4 наведені експериментальні результати 8-кратної дискретизації карт глибини для набору даних «Dolls» порівняно з іншими методами. З виділених областей видно, що методи-аналоги сильно страждають від розмиття країв глибини і артефактів копіювання текстури. Порівняно з наведеними вище методами, запропонований метод дає найкращі результати надроздільної реконструкції карти глибини.

**Деградація за рахунок зменшення вибірки з додаванням шуму.** В реальних умовах, карти глибини, отримані сенсорами, супроводжуються неминучим шумом. Для моделювання таких випадків було проведено тести на наборах даних, представлених в роботі [107], які вводять гаусівський шум з нульовим середнім значенням, дисперсією 25 до наборів даних з пониженою вибіркою при чотирьох коефіцієнтах апсемплінгу. В таблиці 3.5 наведені результати покращення карти глибини за допомогою запропонованого методу, а також шістьма методами аналогами.

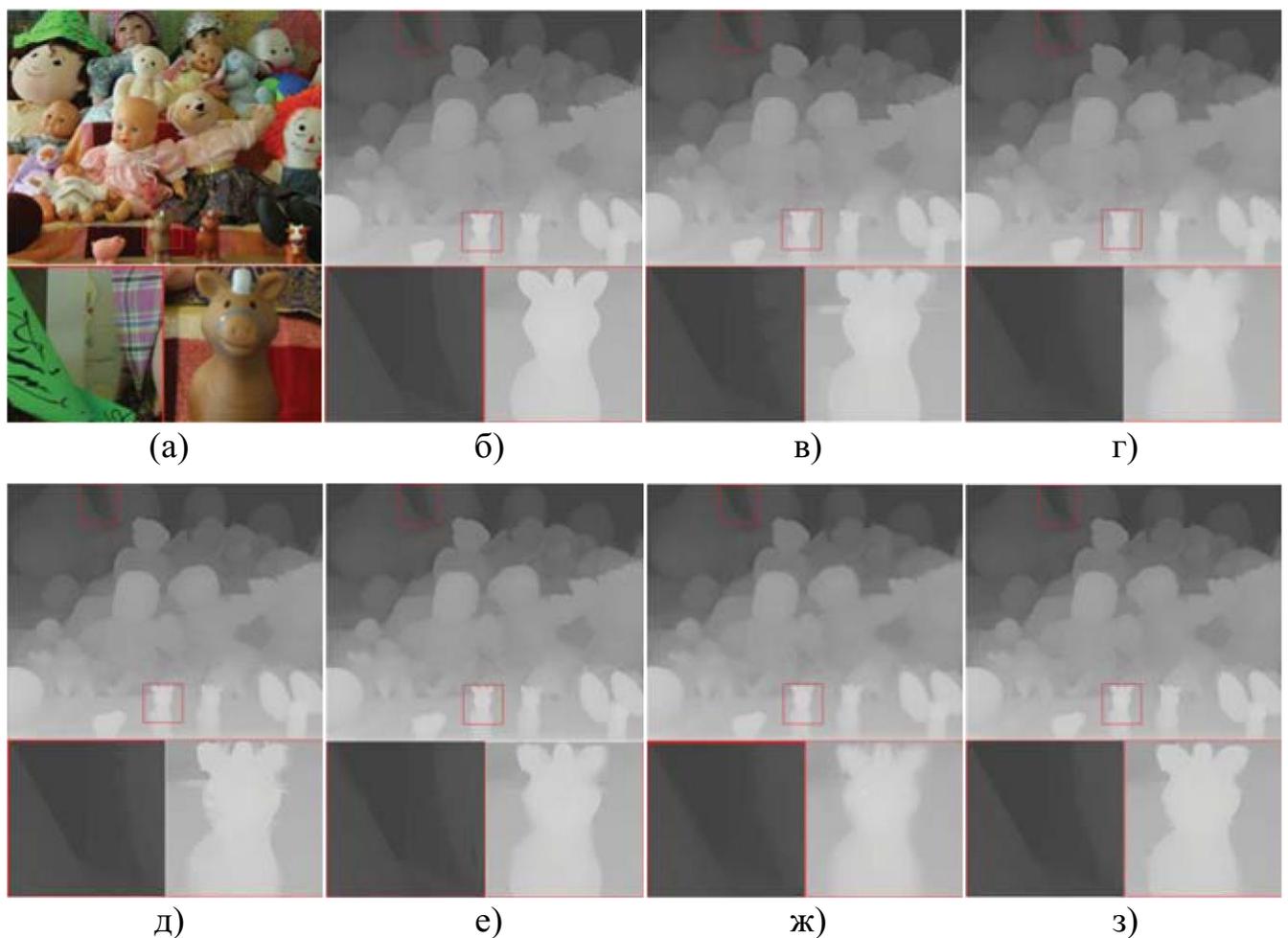


Рисунок 3.4 – Порівняння візуальної якості для надроздільної реконструкції карти глибини на наборі даних «Dolls»: (а) кольорове зображення, (б) істинна карта глибин, карти глибин з дискретизацією (8×) за допомогою (в) Guided, (г) JBU, (д) TGV, (е) Bicubic, (ж) MLS, (з) запропонованого методу

З результатів експерименту видно, що запропонований метод дає найнижчий показник середнього абсолютного відхилення для всіх випадків. Для наочного порівняння результатів на рисунку 3.5 показано результати надроздільної реконструкції карти глибини з шумом, отримані за допомогою запропонованого методу та методів-аналогів. Показано, що методи Guided та MLS залишають сильні шуми, TGV забезпечує чистіші карти глибини, але не зберігає малі структури, такі як палички в чашці. Загалом, запропонований метод здатний подавляти шум і зберегти більшість деталей.

Таблиця 3.5 – Результати роботи методів на вибірці Middlebury, із шумом

Метод	Набір даних															
	Book				Moebius				Dolls				Reindeer			
	2x	4x	8x	16x	2x	4x	8x	16x	2x	4x	8x	16x	2x	4x	8x	16x
Guided	2,09	2,27	2,51	3,25	2,09	2,25	2,57	3,29	2,10	2,23	2,51	3,26	1,30	2,00	3,00	4,15
JBU	0,82	1,40	2,69	3,22	1,19	1,91	2,78	3,56	1,20	1,95	2,81	3,51	1,20	1,60	2,42	3,99
TGV	1,13	1,45	1,82	2,60	1,14	1,46	1,94	2,92	1,15	1,55	2,08	3,03	0,92	1,50	2,87	3,54
Bicubic	0,80	1,22	1,95	3,04	0,56	0,90	1,73	4,01	0,66	1,64	1,76	3,72	0,59	0,84	1,75	4,40
MLS	0,50	0,74	1,21	2,10	0,49	0,75	1,32	2,07	0,58	0,94	1,40	2,15	0,56	0,87	1,40	2,30
Власний	0,44	0,65	1,08	1,81	0,45	0,67	1,17	1,84	0,52	0,85	1,26	1,93	0,53	0,82	1,31	2,14

Однак, в результаті застосування запропонованого методу на наборі даних «Моебіус» спостерігаються деякі розмиті артефакти біля малої частини краю. Основна причина полягає в тому, що цей слабкий край на кольоровому зображенні не може бути виявлений детектором Кенні з попередньо встановленими порогамі.

Навпаки, відповідна область на еталонній карті глибини має сильний розрив глибини. Відповідно до запропонованого методу, це є випадок справжньої неузгодженості і кольорове зображення не приймається як орієнтир для апсемплінгу. Тому це може призвести до артефактів розмиття у випадку високих коефіцієнтів дискретизації, наприклад, 8×, 16×.

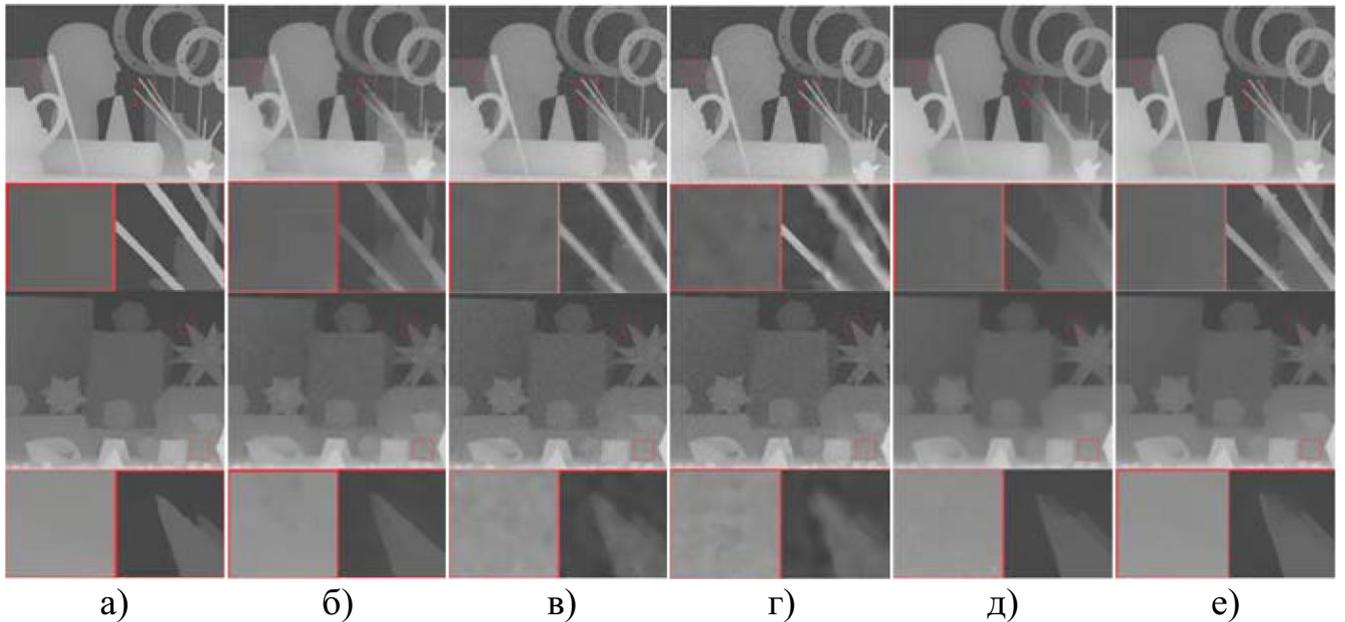


Рисунок 3.5 – Порівняння візуальної якості для надроздільної реконструкції карти глибини: (а) істинна карта глибин, карти глибин з дискретизацією ( $8\times$ ) за допомогою (б) Guided, (в) JBU, (г) TGV, (д) Bicubic, (е) запропонованого методу

#### Деградація в наслідок структурних помилок та випадкових пропусків.

Для кількісної перевірки ефективності побудови карти глибини за допомогою запропонованого методу, був використаний набір даних, представлений в роботі [107], в якому були вручну додані деякі прогалини в еталонні дані бенчмарку Middlebury. Прогалини складаються зі структурних помилок та випадкових пропусків, які генеруються біля країв глибини та в гладких областях відповідно. В таблиці 3.6 наведені результати проведеного експерименту.

Таблиця 3.6 – Результати роботи методів на синтетичному наборі даних

Метод	Набір даних			
	Book	Moebius	Dolls	Reindeer
Guided	0,61	0,66	0,76	0,95
JBU	0,58	0,72	0,82	0,79
TGV	0,63	0,70	0,77	0,92
Bicubic	0,63	0,67	0,76	0,96
MLS	0,60	0,60	0,73	0,76
Власний	0,52	0,56	0,68	0,70

Результати показують, що запропонований метод має найнижчий показник середнього абсолютного відхилення на всіх наборах даних. Візуальні результати роботи запропонованого методу та методів-аналогів наведені на рисунку 3.6. Хоча всі методи дають задовільні результати при добудові карти глибини, результати запропонованого методу можуть забезпечити більш точні межі глибини, що показано у виділених регіонах.

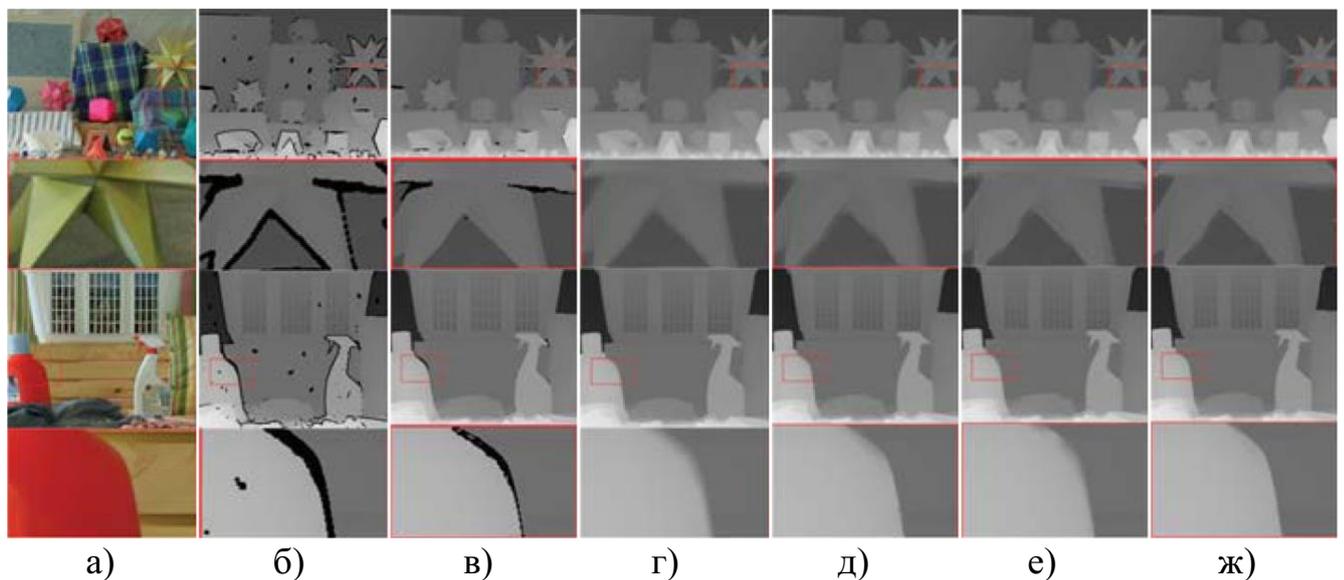


Рисунок 3.6 – Порівняння візуальної якості заповнення карти глибин на наборах даних із структурними помилками та випадковими пропусками: (а) кольорові зображення, (б) погіршені карти глибин, (в) істинні карти глибини, карти глибини, отримані за допомогою (г) Guided, (д) JBU, (е) MLS та (ж) запропонованого методу

### 3.3.5.2 Експерименти з удосконалення карти глибини з використанням реальних наборів даних

Запропонований метод також було протестовано на наборах даних ToF-Mark [84] та NYU [106], дані яких відповідають даним, які можна отримати за допомогою двох типів датчиків глибини відповідно (тобто, датчика глибини ToF та датчика глибини Structured light). Метою експерименту є демонстрація того, що

запропонований метод реконструювати високоякісні карти глибини на базі низькоякісних карт глибини, отриманих різними типами датчиків.

**Експерименти на наборах даних ToF-Mark.** Запропонований метод оцінювався на наборах даних ToF-Mark, що складаються з трьох наборів RGB-D даних: «Book», «Shark» та «Devil» з еталонними картами глибини. Роздільна здатність вихідних карт глибин  $120 \times 160$ , а відповідні зображення інтенсивності мають розмір  $610 \times 810$ . Розрахований коефіцієнт дискретизації становить приблизно  $6,25 \times$  [95]. Таблиця 3.7 ілюструє результати кількісного порівняння запропонованого методу з методами-аналогами. Похибки апсемплінгу обчислені за допомогою середнього абсолютного відхилення в міліметрах.

Таблиця 3.7 – Результати роботи методів на наборі даних ToF-Mark

Метод	Набір даних		
	Books	Shark	Devil
Guided	16,24	17,79	16,68
JBU	14,32	15,89	15,37
TGV	14,49	16,26	14,98
Bicubic	14,51	16,63	24,98
MLS	12,76	14,82	14,77
Власний	12,23	14,15	13,72

Результати експерименту показують, що запропонований метод дозволяє отримати найнижчу величину середньої абсолютної похибки для всіх наборів даних, порівняно з розглянутими методами-аналогами. На рисунку 3.7 проілюстровані результати візуального покращення карти глибини запропонованого методу та методів-аналогів.

Помічено, що результати методів MLS і JBU все ще містять значну кількість шуму через обмежену здатність до знешумлення, тоді як глибина, покращена за допомогою методів TGV, Guided і запропонованого методу, є набагато кращою. Однак результати TGV і Guided вносять артефакти копіювання текстури в деякі гладкі області (наприклад, око риби в наборі даних «Shark», виділене червоним квадратом). Результати запропонованого методу не мають таких артефактів

копіювання текстури. Крім того, край прямокутної коробки в наборі даних «Shark», виділений червоним квадратом, в результаті застосування запропонованого методу є більш точним, ніж при застосуванні інших методів, що доводить, що запропонований метод може ефективно зберігати краї глибини.

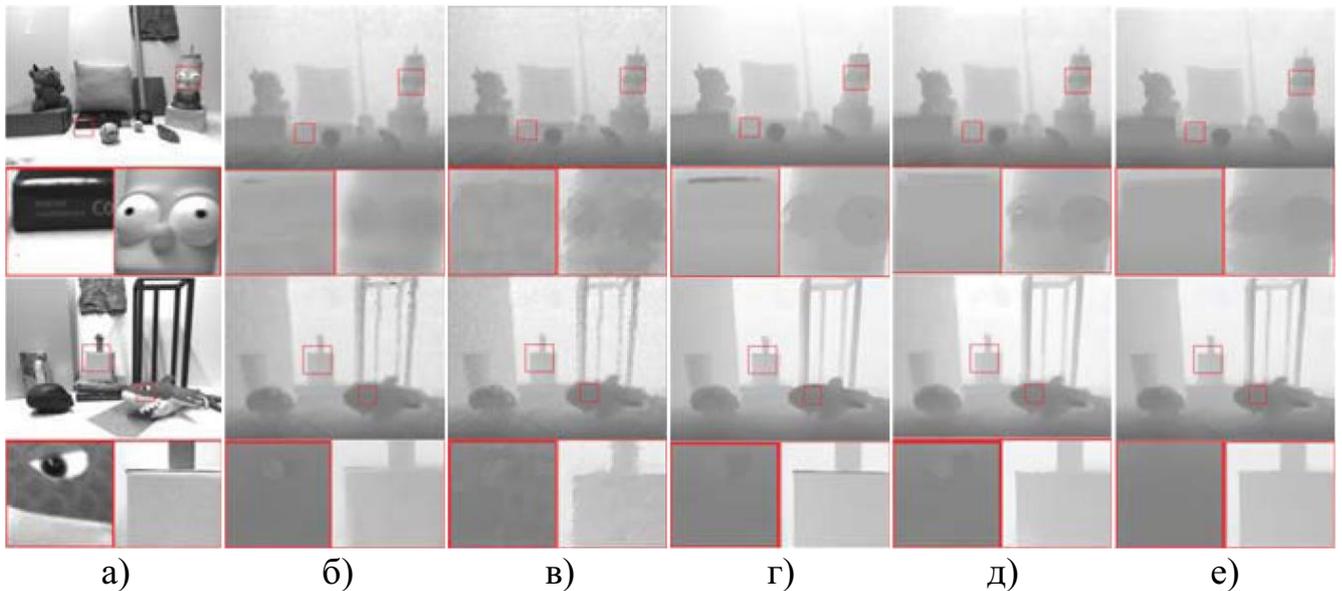


Рисунок 3.7 – Порівняння візуальної якості для карти глибини SR на наборах даних «Devil» та «Shark»: (а) кольорові зображення, карти глибин побудовані за допомогою (б) Guided, (в) JBU, (г) TGV, (д) MLS та (е) запропонованого методу

**Експерименти на наборах даних NYU.** Набір NYU [106] містить дані, зняті датчиками глибини зі структурованим світлом. На рисунку 3.8 наведені результати роботи запропонованого методу порівняно з методами-аналогами.

З виділених областей видно, що розглянуті методи-аналоги (наприклад, MLS та JBU) страждають від артефактів копіювання текстури. На противагу цьому, в результатах запропонованого методу такі артефакти відсутні.

З точки зору збереження країв глибини, метод JBU показує найкращі результати серед існуючих методів, але поступається запропонованому методу. Таким чином, запропонований метод може забезпечити більш надійні результати заповнення карти глибини в реальних сценах, ніж розглянуті методи-аналоги.

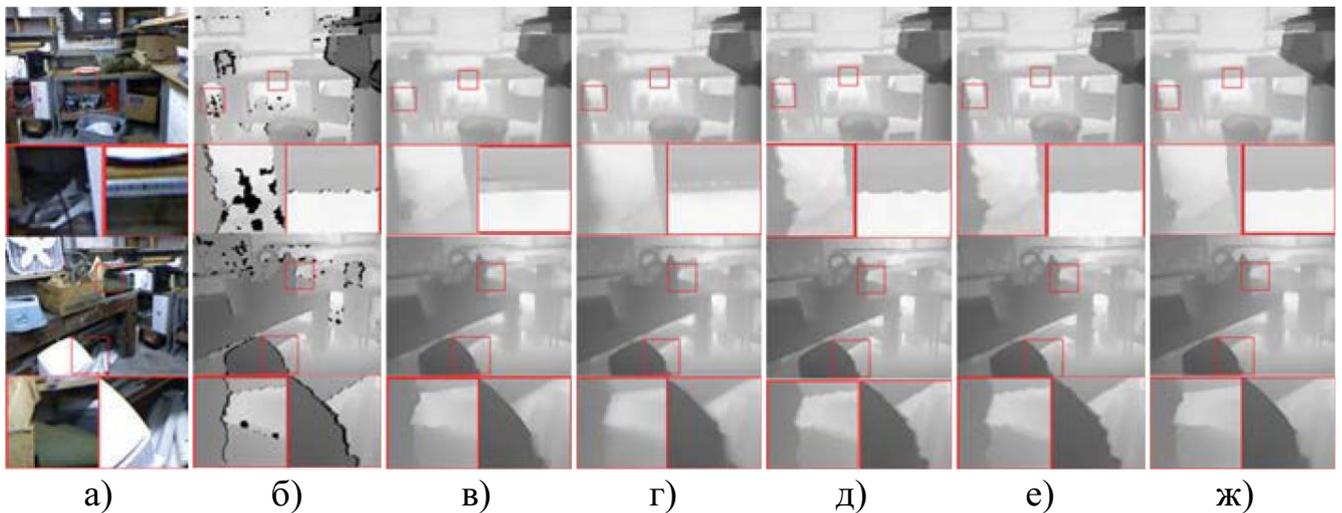


Рисунок 3.8 – Порівняння візуальної якості добудови карти глибин для набору даних NYU: (а) кольорові зображення, (б) зареєстровані початкові карти глибини з Kinect v1, карти глибини добудовані (в) Guided, (г) MLS, (д) JBU, (е) TGV та (ж) запропонованим методом

### 3.3.5.3 Експерименти з вирішення завдань надроздільної реконструкції карти глибини та добудови карти глибини

У попередніх експериментах було показано ефективність запропонованого методу для надроздільної реконструкції карти глибини та добудови карти глибини незалежно один від одного. Для подальшої перевірки робастності запропонованого методу, був проведений експеримент, який стосується надзвичайно складного випадку, в якому запропонований метод має підвищити дискретизацію карти глибин низької роздільної здатності (коефіцієнт 4x) та заповнити прогалини одночасно. В експерименті використано набори даних Нью-Йоркського університету [106]. На рисунку 3.9 наведені результати запропонованого методу в порівнянні з методами-аналогами.

З виділених областей видно, що запропонований метод забезпечує найкращі показники заповнення прогалин, зменшення артефактів копіювання текстури та збереження країв глибини. На противагу цьому, методи-аналоги показують артефакти копіювання текстури (наприклад, виділені в другому ряду) і розмиття

глибинних країв (наприклад, виділені в четвертому ряду) в таких результатах. Деякі з методів-аналогів не можуть дати задовільних результатів у випадку незаповнених прогалів.

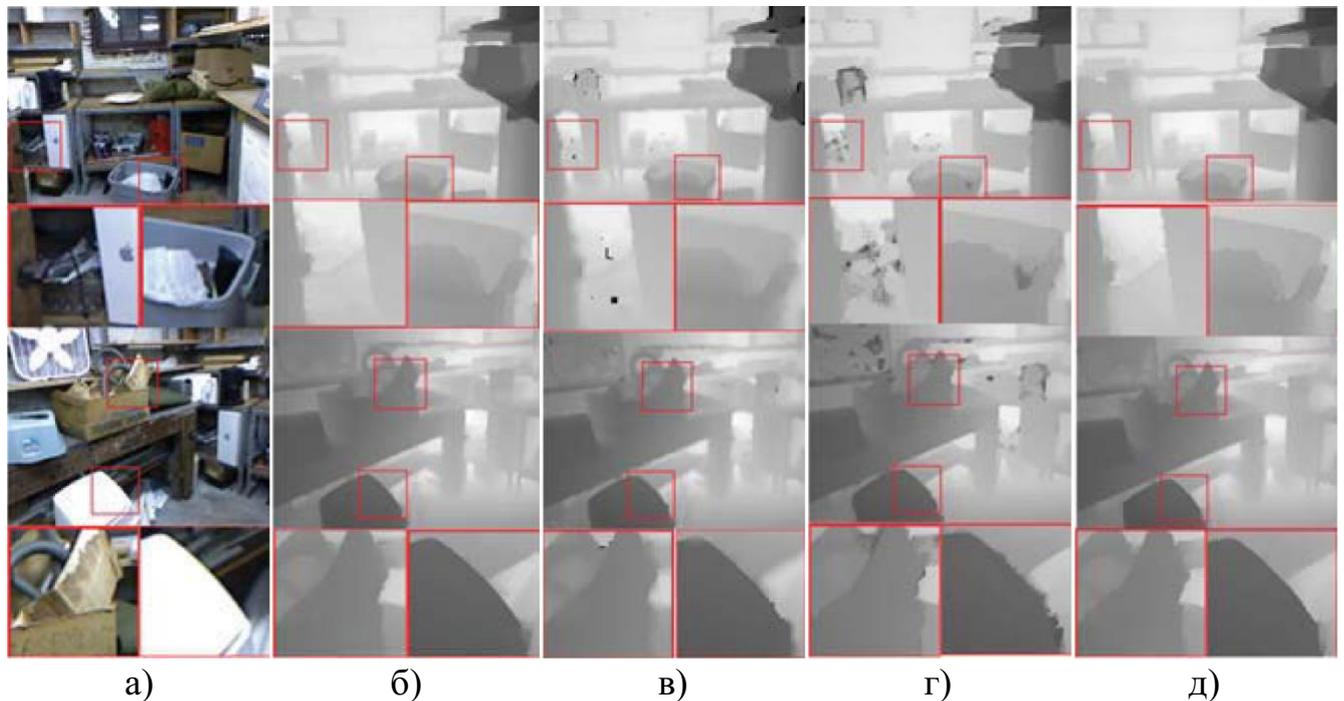


Рисунок 3.9 – Порівняння візуальної якості для покращення карти глибин з комплексною деградацією (зменшення вибірки та відсутність значень глибини) на наборах даних NYU: (а) кольорові зображення з відповідними картами глибини низької роздільної здатності, показаними у верхньому лівому куті, карти глибини покращені за допомогою (б) Guided, (в) JBU, (г) TGV та (д) запропонованого методу.

### **3.4 Удосконалення карти глибини за допомогою використання мінімального кістякового лісу та оцінки неузгодженості країв з м'яким прийняттям рішень**

Для кращого збереження структури на карті глибини в цьому дисертаційному дослідженні запропоновано структурну схему для обчислення побідності в термі регуляризації. В рамках дисертаційного дослідження було висунуто гіпотезу, що подібність обчислюється точніше в просторі, який складається з декількох

мінімальних кістякових дерев (така множина називається мінімальним кістяковим лісом).

Подібність обчислюється на основі шляхів, які проходять всередині мінімального кістякового дерева або через сусідні дерева. У запропонованому методі, такими шляхами моделюється локальна структура глибини. Крім того, обчислення неузгодженості з м'яким прийняттям рішення, запропоноване в попередньому розділі, вбудовується у ваги ребер всередині кожного мінімального кістякового дерева, щоб значно зменшити артефакти копіювання текстури.

### 3.4.1 Модифікація функції енергії мінімального кістякового лісу

В даному методі, терм даних визначається на основі М-оцінки Губера [108], яка є компромісом між L1-нормою та L2-нормою:

$$H_T(M) = \|M_{sub} - V\|_{Hb}, \quad (3.19)$$

$$\|X\|_{Hb} = \begin{cases} \sum_i \frac{x_i^2}{2\eta}, & |x_i| \leq \eta \\ \sum_i \eta \left(|x_i| - \frac{\eta}{2}\right), & |x_i| > \eta \end{cases}, \quad (3.20)$$

де  $V$  – множина спостережуваних значень глибини, які фіксуються безпосередньо сенсорами,  $M_{sub}$  є підмножиною  $M$ , яка складається з глибин пікселів, які мають спостережувані значення.

Крім того, порівняно з рівнянням (3.3), в якому подібність  $\xi_f^{kl}$  обчислюється безпосередньо за різницею глибини і кольору, в даному випадку  $\xi_f^{kl}$  обчислюється на основі відстані в мінімальному кістяковому лісі, що показано у вигляді рівняння:

$$\xi_f^{kl} = e^{\frac{-dst(k,l)}{\delta}}, \quad (3.21)$$

де  $\delta$  контролює смугу експоненціальної функції, а  $dst(k, l)$  позначає відстань між пікселями  $k$  та  $l$  в мінімальному кістяковому лісі.

### 3.4.2 Принцип роботи запропонованого методу

Існуючі методи уточнення карти глибини [94, 96, 100], включаючи методи запропоновані в попередніх підрозділах, обчислюють спорідненість члена регуляризації за неструктурною схемою, яка базується лише на різниці кольору та глибини між пікселем та його сусідніми пікселями. Така схема ігнорує локальну структуру на карті глибини. Тому вона може надмірно згладжувати краї глибини на покращеній карті глибини (наприклад,  $16\times$  для глибини SR). На противагу цьому, в роботах [108, 109] було показано, що деревовидний фільтр може забезпечити більш надійну продуктивність у збереженні країв. Він був використаний у багатьох задачах комп'ютерного зору, наприклад, згладжування зі збереженням структури [110], стерео зіставлення [111] тощо. У цих методах мінімальне кістякове дерево використовується для автоматичного видалення двох несхожих пікселів, які знаходяться близько один до одного в просторовій області, що робить відстань між деревами метрикою, яка враховує краї. Однак, використання одного глобального мінімального кістякового дерева (коли одне мінімальне кістякове дерево з'єднує всі пікселі зображення разом і агрегує ефекти згладжування глобально вздовж дерева) стикається з проблемою «протікання» країв. При такій конфігурації, з'єднання неминуче перетинатимуть деякі сильні краї, і ці краї будуть спотворені.

Запропонований метод обчислює подібність  $\xi_s^{kl}$  в термі регуляризації на основі відстані між  $k$  і  $l$  в дереві. А щоб уникнути побудови мінімального кістякового дерева через краї глибини, що призводить до артефактів протікання країв, запропонований метод будує мінімальне кістякове дерево для кожного суперпікселя на кольоровому зображенні, згенерованому надлишковою сегментацією. Оскільки кольорове зображення в нашому випадку є високоякісним,

а пікселі всередині одного суперпікселя мають схожі характеристики, розрахунок відстані у дереві є більш надійним в області суперпікселя. Крім того, через надмірну сегментацію піксель та його сусідні пікселі, які знаходяться на сусідніх мінімальних кістякових деревах, можуть мати схожі значення глибини. Тому для врахування таких ситуацій слід також обчислювати зв'язки країв між сусідніми деревами. Сусідні мінімальні кістякові дерева об'єднуються разом і стають так званим мінімальним кістяковим лісом.

Якщо спорідненість просто обчислюється на основі зареєстрованого кольорового зображення, це призводить до артефактів копіювання текстури та розмиття країв глибини через невідповідність країв між картою глибини та зареєстрованим кольоровим зображенням. На основі запропонованої раніше методу удосконалення карти глибини з м'яким прийняттям рішень, який демонструє сильну здатність зменшувати артефакти копіювання текстури, ваги ребер всередині кожного мінімального кістякового дерева у запропонованому лісі обчислюються на основі цього методу.

### 3.4.2.1 Побудова мінімального кістякового лісу

Для надмірної сегментації кольорового зображення використовується алгоритм суперпіксельної сегментації SLIC [112]. Це дозволяє уникнути побудови мінімального кістякового дерева через краї глибини та зменшити артефакти витікання країв [110]. Цей метод сегментації забезпечує збереження основних країв зображення та має лінійну складність. У кожному сегменті генерується 8-зв'язний зважений підграф  $SubG = (N, E, W)$ . Всі пікселі в цьому сегменті є вершинами  $N$ ,  $E$  складається з усіх ребер між вершинами,  $W$  – це множина ваг ребер, яка визначається за допомогою наступного рівняння:

$$W(k, l) = |\Delta_{clr}^{kl}|, \quad (3.22)$$

де  $\Delta_{clr}^{kl}$  – різниця кольорів між  $k$  та  $l$ . Оскільки кольорове зображення та карта глибини мають різні текстурні патерни, така конфігурація призводить до артефактів копіювання текстури та розмиття границь глибини. Запропонований метод явно вбудовує вимірювання неузгодженості країв у побудову мінімального кістякового дерева для вирішення цих проблем.

Отже, всередині кожного мінімального кістякового дерева, наразі, кожна вага ребра між двома сусідніми вершинами – це їхня колірна схожість, як показано у рівнянні (3.18). На основі таких підграфів можна обчислити мінімальне кістякове дерево на кожному суперпікселі, видаливши надлишкові ребра. Через надмірну сегментацію піксель та його сусідні пікселі, які знаходяться на сусідніх деревах, все ще можуть мати однакові значення глибини.

Тому, щоб врахувати цю ситуацію, також мають бути обчислені ребра, що з'єднують сусідні дерева. Щоб узгодити вагу ребер всередині кожного дерева, вагу ребра, що з'єднує сусідні дерева, обчислюють за кольором та різницею глибини, щоб зменшити артефакти копіювання текстури. Зокрема, таке ребро, визначене пікселями з подібним кольором і глибиною, які знаходяться на сусідніх мінімальних кістякових деревах, може розширити дерево на суперпіксель до мінімального кістякового лісу на всьому зображенні. Це можна розглядати як контроль для подолання проблеми «протікання» країв.

Подібне налаштування також використовується в роботі [110], яка враховує лише схожість кольорів для виконання згладжування з урахуванням країв. Насправді, запропонований мінімальний кістяковий ліс можна розглядати як простір відстаней на основі вмісту зображення. Просторовий фактор включено до шляху між пікселями в лісі. Порівняно з неструктурним евклідовим простором відстаней, який використовується у двох окремих ядрах білатерального фільтра [113], відстань, обчислена у запропонованому мінімальному кістяковому лісі, представляє структуру зображення. Запропонований метод моделює кожен піксель  $px$  як двовимірну точку  $Pnt_{px} = (c_{px}, d_{px})$ , яка включає його значення кольору  $c_{px}$  і значення глибини  $d_{px}$ . Метрика відстані між двома пікселями визначається як L1-

норма. Оптимальна пара вузлів  $(k^*, l^*)$  між сусідніми суперпікселями  $SPX_a$  та  $SPX_b$  визначається за допомогою наступного рівняння:

$$(k^*, l^*) = \min_{\substack{k \in SPX_a \\ l \in SPX_b}} \|Pnt_k - Pnt_l\|, \quad (3.23)$$

У випадку використання алгоритму вичерпного пошуку, складність становить  $O(n_1 n_2)$ , де  $n_1, n_2$  – кількість вузлів у їхніх мінімальних кістякових деревах відповідно. Якщо розмір суперпікселя є великим, ця операція може зайняти багато часу. Запропонований метод використовує більш ефективний алгоритм «розділяй і володарюй» [114] для вирішення цієї задачі, що зменшує складність обчислень до  $O((n_1 + n_2) \log(n_1 + n_2))$ .

Отже, відстань між двома сусідніми вузлами в термі регуляризації обчислюється наступним чином: якщо  $k$  та  $l$  розташовані в одному суперпікселі, відстань між ними можна обчислити за допомогою стандартної деревовидної відстані вздовж шляху на мінімальному кістяковому дереві наступним чином:

$$dst(k, l) = \sum_{i=0}^n W(k_i, k_i + 1), \quad (3.24)$$

де  $(k_i, k_i + 1)$  – це сусідні вузли вздовж шляху. Така відстань ефективно обчислюється на основі методу найменшого спільного предка (LCA) [115].

В іншому випадку відстань обчислюється за допомогою трьох компонентів наступним чином:

$$dst(k, l) = dst(k, k') + dst(l, l') + 0,5 \times (|\Delta_{clr}^{k'l'}| + |\Delta_{dpt}^{k'l'}|), \quad (3.25)$$

де  $k, l$  – дві найближчі пари вузлів у сусідніх мінімальних кістякових деревах відповідно. Вони є найближчими з огляду на схожість кольору та глибину (останній

доданок у правій частині рівняння). Значення  $\Delta_{clr}^{k'l'}$  і  $\Delta_{dpt}^{k'l'}$  представляють різницю кольору і глибини між пікселями  $k'$  і  $l'$  на кольоровому зображенні і грубо інтерпольованій карті глибини відповідно. Значення  $(k, k')$  і  $(l, l')$  – це пари пікселів, розташованих в одному мінімальному кістяковому дереві.

### 3.4.2.2 Вбудовування вимірювання неузгодженості країв у мінімальний кістяковий ліс

Представлений вище метод вимірювання неузгодженості країв з м'яким прийняттям рішень має значні можливості у зменшенні артефактів копіювання текстур. В рамках даного розділу, він описується як двонаправлене оцінювання шляхом зміни ролей карти країв кольорового зображення і карти країв глибини. У кожному напрямку найкращі пари пікселів країв визначаються за допомогою МВП-оптимізації. Терм даних містить інформацію про локальну структуру, яка базується на мінімально зваженому двосторонньому зіставленні, а терм регуляризації містить інформацію про глобальну структуру. Вартість оптимального узгодження пікселів країв є виміром неузгодженості країв, яка знаходиться в діапазоні  $[0, 1]$ .

Запропонований в даному розділі метод вбудовує вимірювання неузгодженості країв з м'яким прийняттям рішень в конструкцію мінімального кістякового дерева, щоб зменшити артефакти копіювання текстури. У запропонованому методі  $\lambda$  визначається як набір довірчих значень для пікселів, який обчислюється відповідно до рівняння (3.11). Чим меншим є значення  $\lambda$ , тим більше узгодженість країв. Зокрема, запропонований метод модифікує визначення  $W$  у рівнянні (3.18) для ребра  $(k, l)$  наступним чином:

$$W'(k, l) = |\Delta_{clr}^{kl}| \times (1 - \lambda_{kl}) + |\Delta_{dpt}^{kl}| \times \lambda_{kl} \quad (3.26)$$

де  $\lambda_{kl}$  – довірче значення для пари пікселів  $k$  та  $l$ , яке визначається як  $\lambda_{kl} = \max(\lambda(k), \lambda(l))$  для кращого збереження країв глибини;  $\lambda(k)$  та  $\lambda(l)$  – довірчі

значення для  $k$  та  $l$  відповідно. Коли карта країв кольорового зображення краще узгоджується з картою країв глибини (значення  $\lambda_{kl}$  наближене до нуля), колір  $\Delta_{clr}^{kl}$  може відігравати важливішу роль в обчисленні ваг ребер всередині мінімального кістякового дерева, і навпаки.

В рамках даного методу, формула обчислення ваг ребер (3.22) замінює оригінальні обчислення  $W$  у рівняннях (3.18), (3.20) та (3.21). Порівняно з попередніми методами покращення карти глибини на основі Марківського випадкового поля, які не використовують інформацію про структуру при обчисленні члена регуляризації, запропонований метод може краще зберігати краї глибини, використовуючи метрику відстані у мінімальному кістяковому дереві, визначену в рівняннях (3.20) і (3.21). Крім того, завдяки запропонованій покращеній процедурі побудови мінімального кістякового дерева із вбудовуванням методу вимірювання неузгодженості країв з м'яким прийняттям рішень, в процесі удосконалення карти глибини вдається зменшити появу артефактів копіювання текстури.

### **3.4.2.3 Адаптація пропускної здатності для обчислення спорідненості в Марківському випадковому полі**

Під час дослідження роботи методу було зроблено спостереження, що ширина смуги пропускання  $\delta$  у рівнянні (3.17) може впливати на продуктивність моделі. Наприклад, коли два пікселі, що застосовуються в регуляризаційному члені Марківського випадкового поля, мають явно різні значення глибини, мала смуга пропускання  $\delta$  може забезпечити кращу продуктивність. Тому, якщо відома різниця глибин між кожною сусідньою парою пікселів у термі регуляризації, то смугу пропускання  $\delta$  можна адаптивно вибирати як попередню при реконструкції високоякісної карти глибини. Цей попередній вибір може зберегти краї глибини і зменшити артефакти, спричинені шумом і копіюванням текстур.

У запропонованому методі для оцінки такого попереднього значення використовується групо інтерпольована карта глибини  $\hat{M}$ . Через низьку якість  $\hat{M}$  розташування країв зміщені щодо реальних країв. Тому оцінювати різницю глибин між парою пікселів  $k, l$  шляхом прямого обчислення на них самих є некоректним. Запропонований метод визначає дві підобласті (емпірично, розмір областей був підібраний як  $11 \times 11$ ), які центровані відповідно коло  $k, l$ , щоб сформувати набори пікселів  $S_k$  та  $S_l$ . Максимальна абсолютна різниця між цими наборами  $Df_{max}$  може бути використана для аналізу потенційного попереднього значення для  $k$  та  $l$ . Якщо значення  $Df_{max}$  велике, це вказує на те, що пара пікселів  $k, l$  знаходиться близько до країв глибини і відповідному  $\delta$  має бути присвоєне невелике значення для кращого збереження країв глибини. Навпаки, якщо  $Df_{max}$  мале, це означає, що пара пікселів  $k, l$  знаходиться на гладкій ділянці. У цьому випадку для  $\delta$  слід вибрати велике значення для згладжування шуму і подальшого зменшення артефактів копіювання текстури.

Проста реалізація обчислення  $Df_{max}$  полягає в обчисленні різниць для всіх елементів двох множин  $S_k, S_l$  зі складністю  $O(t^2)$ , де  $t$  – кардинальне число для кожної множини – така реалізація потребує багато часу.

В рамках даного методу запропоновано більш ефективний алгоритм, який має складність  $O(t)$ . Спочатку обчислюються мінімальне та максимальне значення глибини (тобто,  $max_k, min_k, max_l, min_l$ ) для множин  $S_k$  та  $S_l$  відповідно. Значення  $Df_{max}$  обчислюється наступним чином:

$$Df_{max} = \max(|max_k - min_l|, |max_l - min_k|) \quad (3.23)$$

Оскільки ситуації з шумом різноманітні, важко змоделювати функцію відображення між  $\delta$  і  $Df_{max}$  як фіксовану модель (наприклад, лінійну). У запропонованому методі  $Df_{max}$  обчислюється для кожної пари пікселів у термі регуляризації, і всі пари пікселів класифікуються на три категорії з використанням подвійних порогів  $\omega_l$  та  $\omega_h$ . Такі категорії представляють області біля сильних

країв, області біля слабких країв та згладжені області. Для кожного типу областей вибираються різні значення  $\delta$ , які визначаються наступним рівнянням:

$$\delta = \begin{cases} 2, & Df_{max} > \omega_h \text{ (сильний край)} \\ 10, & Df_{max} < \omega_l \text{ (гладка область)} \\ 4, & \text{в іншому випадку (слабкий край)} \end{cases} \quad (3.27)$$

Оптимальні значення  $\delta$  були визначені на наборі даних «Laundry» і застосовані до решти експериментів.

На рисунку 3.10 наведено суб'єктивне порівняння покращених карт глибини між неструктурною схемою, вбудованою з моделлю неузгодженості [116], запропонованим методом без підбору  $\delta$  та запропонованим методом з підбором  $\delta$  на безшумних наборах даних Middlebury. При використанні запропонованого методу з підбором  $\delta$  можна спостерігати покращення результатів.

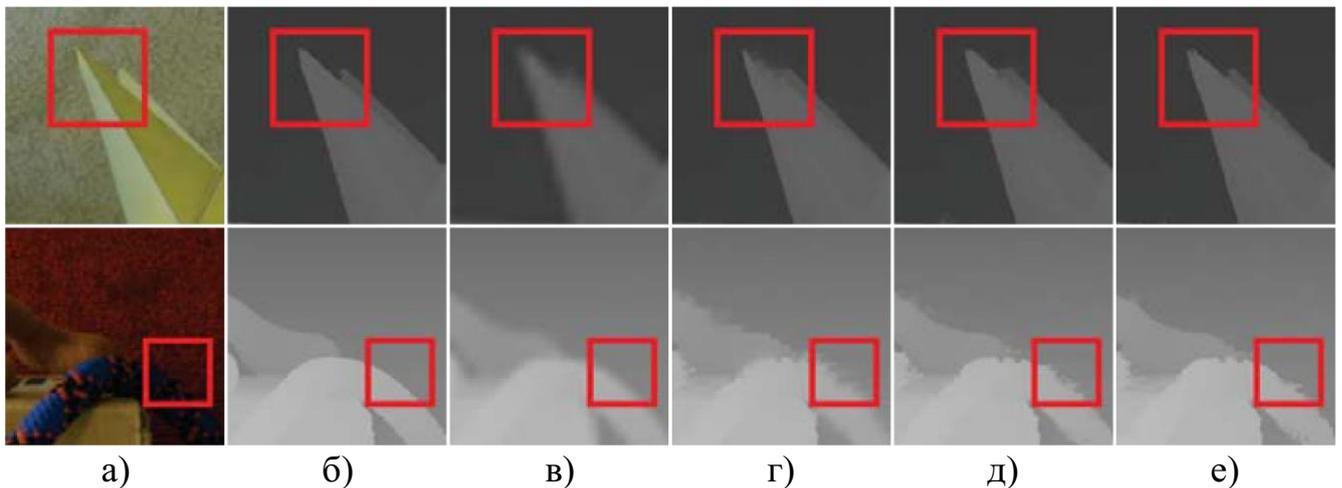


Рисунок 3.10 – Візуальне порівняння дискретизованих ( $8\times$ ) карт глибини, (а) кольорові ділянки, (б) істинні карти глибини, карти глибини низької роздільної дискретизовані: (в) Bicubic, (г) лише метод неузгодженості, (д) МКЛ + метод неузгодженості, (е) МКЛ + метод неузгодженості + адаптація параметрів

### 3.4.3 Експериментальне дослідження запропонованого методу

Ефективність запропонованого методу оцінено для надроздільної реконструкції карти глибин та побудови карти глибин на різних наборах даних.

Під час експериментального дослідження запропонованого методу були проведені наступні експерименти:

- 1) дослідження впливу значень параметрів методу на його ефективність;
- 2) дослідження роботи методу в процесі надроздільної реконструкції карти глибини на синтетичних наборах даних та на реальних даних;
- 3) дослідження роботи методу в процесі побудови карти глибини на синтетичних та реальних наборах даних;
- 4) дослідження середнього часу роботи методу.

#### 3.4.3.1 Дослідження впливу значень параметрів методу на його ефективність

На першому етапі експериментального дослідження роботи запропонованого методу було проведено дослідження впливу деяких ключових параметрів, включаючи кількість суперпікселів та коефіцієнт балансу  $\xi$  у рівнянні (3.3).

**Дослідження кількості суперпікселів.** В даному експерименті змінювалася кількість суперпікселів, тоді як інші параметри залишались незмінними ( $\eta = 10$  у рівнянні (3.16),  $\omega_l = 5$ ,  $\omega_h = 30$  та  $\delta = [2,10,4]$  для відповідних випадків у рівнянні (3.25)). В процесі дослідження був алгоритм Бройдена – Флетчера – Гольдфарба – Шанно з обмеженим використанням пам'яті, щоб запам'ятати 10 попередніх кроків оптимізації для побудови наближеної матриці Гессе. На рисунку 3.12 наведено ілюстрацію результату надроздільної реконструкції карти глибини для набору даних «Laundry» у випадку коефіцієнтів дискретизації  $2\times$ ,  $4\times$ ,  $8\times$  та  $16\times$ . Можна помітити, що ефекти, спричинені різною кількістю суперпікселів, є незначними, особливо для малих коефіцієнтів дискретизації (наприклад,  $2\times$ ,  $4\times$ ). Оптимальне значення становить близько 700, що забезпечує надійні результати для всіх

випадків у цьому експерименті. У наступних експериментах воно фіксується на рівні 700.

**Дослідження значення коефіцієнту балансу  $\xi$ .** Коефіцієнт балансу  $\xi$  у рівнянні (3.3) можна проаналізувати аналогічно до того, як це було зроблено в розділі 3.3.5. В запропонованому методі він дорівнює 0,01 для всіх коефіцієнтів дискретизації при дискретизації карт глибин без додавання шуму. Для надроздільної реконструкції карти глибини з додаванням шуму,  $\xi$  встановлюється рівним 5, 3,3, 2,5 і 1,43 для  $2\times$ ,  $4\times$ ,  $8\times$  і  $16\times$  відповідно. Для побудови карти глибини  $\xi$  дорівнює 0,1 та 10 для всіх експериментів з заповнення карти глибини на синтетичних наборах даних та реальних наборах даних відповідно. Слід зазначити, що запропонований метод застосовує значення  $\xi$ , скориговане на основі набору даних «Laundry» для надроздільної реконструкції карти глибини та побудови незашумленої карти глибини. Для побудови зашумленої карти глибин на наборах даних NYU, для налаштування  $\xi$  використовується лише одна пара зображень.

#### **3.4.3.2 Дослідження роботи методу в процесі надроздільної реконструкції карти глибини**

В даному дослідженні оцінюється робота методу в процесі надроздільної реконструкції карти глибини з використанням синтетичних наборів даних та реальних даних.

**Надроздільна реконструкція карти глибини на синтетичних наборах даних.** У цій частині наведено експерименти на синтетичних наборах даних, які включають два види деградації: субдискретизація без шуму та з додаванням шуму.

*Деградація шляхом субдискретизації без додавання шуму.* В даному випадку експерименти проводилися на чистих від шуму наборах даних Middlebury, зменшеними за допомогою інтерполяції найближчого сусіда. Запропонований метод було протестовано для кількох коефіцієнтів наддискретизації ( $2\times$ ,  $4\times$ ,  $8\times$ ,  $16\times$ ) і порівняно з методами-аналогами. В таблиці 3.8 наведені результати наддискретизації з чотирма різними коефіцієнтами.

Таблиця 3.8 – Результати роботи методів на вибірці Middlebury, без шуму

Метод	Набір даних															
	Book				Moebius				Dolls				Reindeer			
	2x	4x	8x	16x	2x	4x	8x	16x	2x	4x	8x	16x	2x	4x	8x	16x
Guided	0,23	0,36	0,59	1,15	0,24	0,39	0,61	1,17	0,29	0,36	0,57	1,15	0,43	0,56	0,89	1,82
JBU	0,18	0,37	0,75	1,58	0,19	0,38	0,77	1,48	0,22	0,40	0,76	1,48	0,28	0,52	1,02	1,90
TGV	0,20	0,28	0,43	0,84	0,21	0,30	0,51	0,89	0,23	0,39	0,72	2,21	0,33	0,51	1,05	3,07
Bicubic	0,16	0,30	0,61	1,16	0,14	0,32	0,61	1,15	0,22	0,40	0,68	1,20	0,32	0,57	1,01	1,89
MLS	0,17	0,28	0,47	1,18	0,17	0,36	0,51	0,94	0,25	0,40	0,62	0,99	0,34	0,65	0,77	1,45
Власн-Ж	0,22	0,31	0,59	1,24	0,18	0,37	0,59	1,09	0,27	0,40	0,74	1,43	0,38	0,59	0,99	2,35
Власн-М	0,17	0,30	0,48	1,19	0,15	0,33	0,53	0,90	0,24	0,38	0,69	1,38	0,35	0,53	0,84	1,75
Власн-МКЛ	0,15	0,25	0,42	1,06	0,12	0,28	0,46	0,78	0,20	0,32	0,52	0,90	0,31	0,58	0,69	1,30

Загалом, запропонований метод дає найнижче значення середнього абсолютного відхилення для більшості випадків. На рисунку 3.11 наведені експериментальні результати 8-кратної наддискретизації карт глибини для наборів даних «Laundry» та «Dolls» у порівнянні з методами-аналогами.

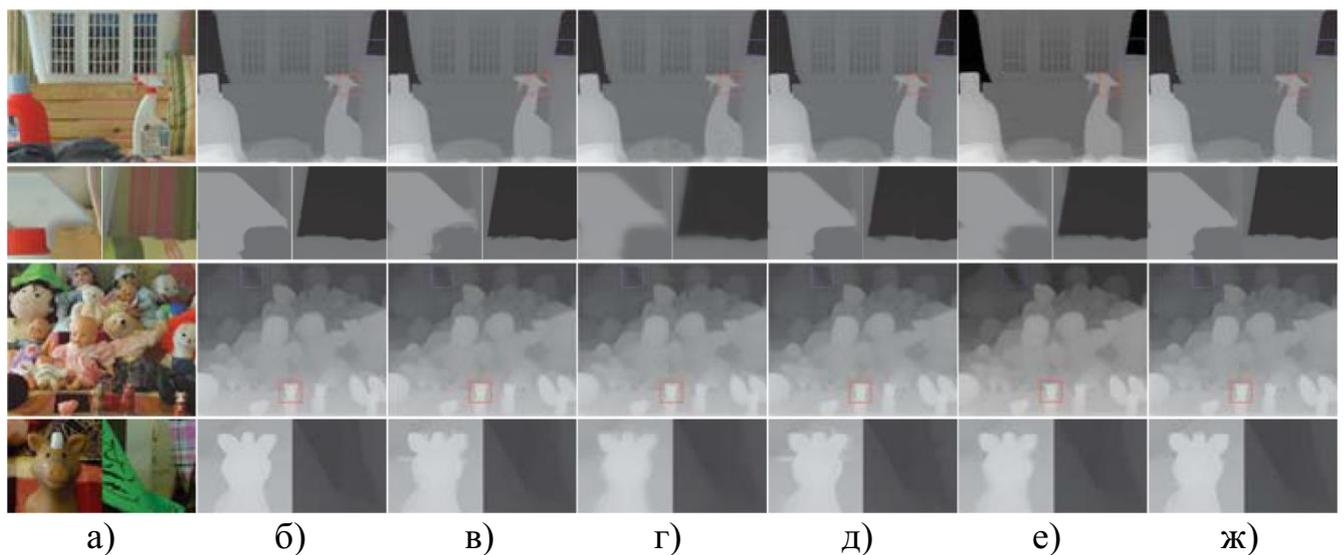


Рисунок 3.11 – Візуальне порівняння дискретизованих (8×) карт глибини на безшумних наборах даних Middlebury: (а) кольорові зображення високої роздільної здатності, (б) істинні карти глибини, карти глибини низької роздільної здатності, дискретизовані за допомогою: (в) JBU, (г) TGV, (д) Guided, (е) MLS і (ж) запропонованого методу

Можна помітити, що метод TGV сильно постраждав від артефактів копіювання текстури. Крім того, з виділених областей деякі артефакти біля країв глибини з'являються на результатах інших методів. Результати запропонованого методу є найбільш близькими до еталонних карт глибин без артефактів копіювання текстури та розмиття країв глибин.

*Деградація шляхом субдескриптивізації з додаванням шуму.* В реальних ситуаціях, карти глибини, отримані сенсорами, завжди будуть зашумлені. Для моделювання таких випадків слід проводити дослідження на зашумлених наборах даних [107]. В таблиці 3.9 наведені результати покращення карти глибини за допомогою запропонованого методу, а також методами-аналогами. З таблиці видно, що запропонований метод дає найнижчий показник середнього абсолютного відхилення для всіх випадків. На рисунку 3.12 наведені результати надроздільної реконструкції карти глибини на зашумлених наборах даних «Book» та «Moebius». Карты глибин низької роздільної здатності побудовано за допомогою методів-аналогів та за допомогою запропонованого методу. На рисунку видно, що методи Guided і JBU залишають сильні шуми. Хоча TGV і MLS можуть забезпечити чистіші результати, вони сильно страждають від артефактів копіювання текстури.

Таблиця 3.9 – Результати роботи методів на вибірці Middlebury, із шумом

Метод	Набір даних															
	Book				Moebius				Dolls				Reindeer			
	2x	4x	8x	16x	2x	4x	8x	16x	2x	4x	8x	16x	2x	4x	8x	16x
Guided	2,09	2,27	2,51	3,25	2,09	2,25	2,57	3,29	2,10	2,23	2,51	3,26	1,30	2,00	3,00	4,15
JBU	0,82	1,40	2,69	3,22	1,19	1,91	2,78	3,56	1,20	1,95	2,81	3,51	1,20	1,60	2,42	3,99
TGV	1,13	1,45	1,82	2,60	1,14	1,46	1,94	2,92	1,15	1,55	2,08	3,03	0,92	1,50	2,87	3,54
Bicubic	0,80	1,22	1,95	3,04	0,56	0,90	1,73	4,01	0,66	1,64	1,76	3,72	0,59	0,84	1,75	4,40
MLS	0,50	0,74	1,21	2,10	0,49	0,75	1,32	2,07	0,58	0,94	1,40	2,15	0,56	0,87	1,40	2,30
Власний	0,42	0,63	1,05	1,78	0,44	0,65	1,15	1,82	0,49	0,82	1,23	1,89	0,52	0,77	1,28	2,12

Порівняно з цими методами, використання запропонованого методу дозволяє значно подавити шум, а також отримати більш точні краї без артефактів копіювання текстури.

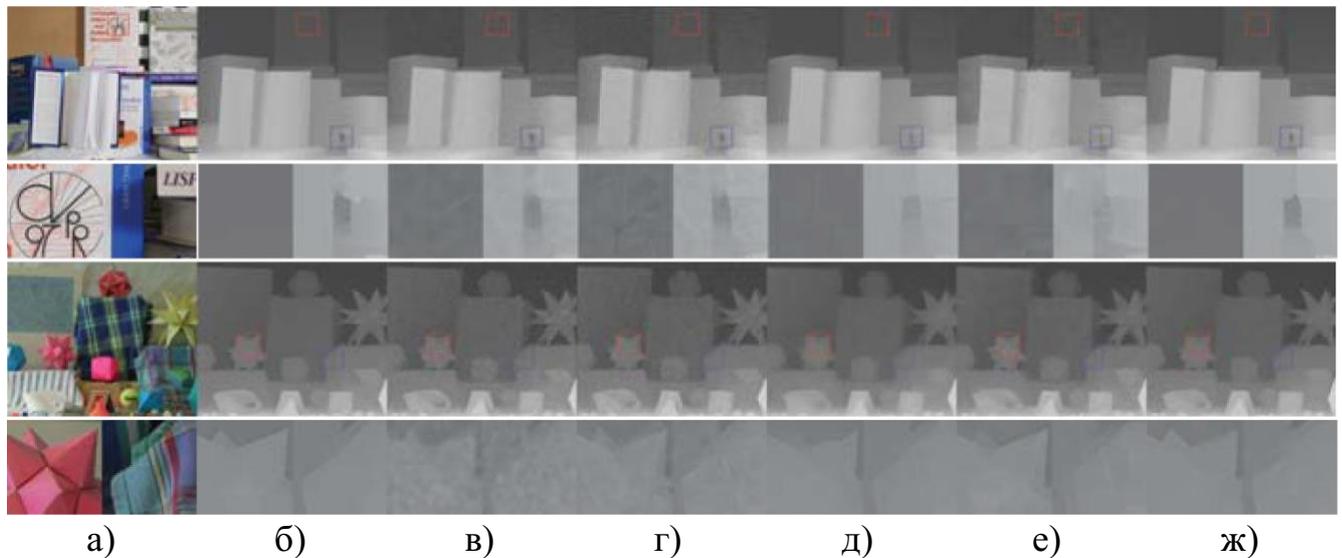


Рисунок 3.12 – Візуальне порівняння дискретизованих ( $8\times$ ) карт глибин на зашумлених наборах даних Middlebury: (а) кольорові зображення високої роздільної здатності, (б) істинні карти глибини, карти глибини низької роздільної здатності, дискретизовані за допомогою: (в) JBU, (г) TGV, (д) Guided, (е) MLS і (ж) запропонованого методу

**Надроздільна реконструкція карти глибини на реальних даних.** Запропонований метод було протестовано на наборах даних ToF-Mark [84] для надроздільної реконструкції карти глибини, щоб довести робастність методу на реальних наборах даних, отриманих за допомогою ToF-сенсорів.

Роздільна здатність карт глибини становить  $120\times 160$ , а зареєстровані зображення інтенсивності мають розмір  $610\times 810$ . Запропонований коефіцієнт дискретизації становить приблизно  $6,25\times$ .

В таблиці 3.10 наведені результати кількісного порівняння запропонованого методу та методів-аналогів. Похибки апроксимації розраховані за допомогою метрики середнього абсолютного відхилення в міліметрах.

Таблиця 3.10 – Результати роботи методів на наборі даних ToF-Mark

Метод	Набір даних		
	Books	Shark	Devil
Guided	16,24	17,79	16,68
JBU	14,32	15,89	15,37
TGV	14,49	16,26	14,98
Bicubic	14,51	16,63	24,98
MLS	12,76	14,82	14,77
Власний	12,23	14,15	13,72
Власний-МКЛ	11,81	13,91	13,50

Результати запропонованого методу показують найнижчу похибку середнього абсолютного відхилення для всіх трьох наборів даних порівняно з методами-аналогами.

На рисунку 3.15 наведені візуальні результати запропонованого методу порівняно з методами-аналогами. Загалом, результати MLS та JBU все ще містять значну кількість шуму, що свідчить про їхню обмежену здатність до знешумлення. Карти глибин, отримані за допомогою TGV, MLS та запропонованого методу, є чистішими.

Однак, на виділених ділянках можна спостерігати артефакти копіювання текстури (наприклад, другий рядок на рисунку 3.13) та розмиття країв глибини (наприклад, п'ятий рядок на рисунку 3.13) у результатах TGV та Bicubic. Результати запропонованого методу не мають таких артефактів. Крім того, край прямокутної рамки в наборі даних «Shark», виділений червоним квадратом (четвертий рядок на рисунку 3.13), є більш різким у результатах запропонованого методу, що свідчить про те, що запропонований метод дозволяє ефективно зберігати краї глибини.

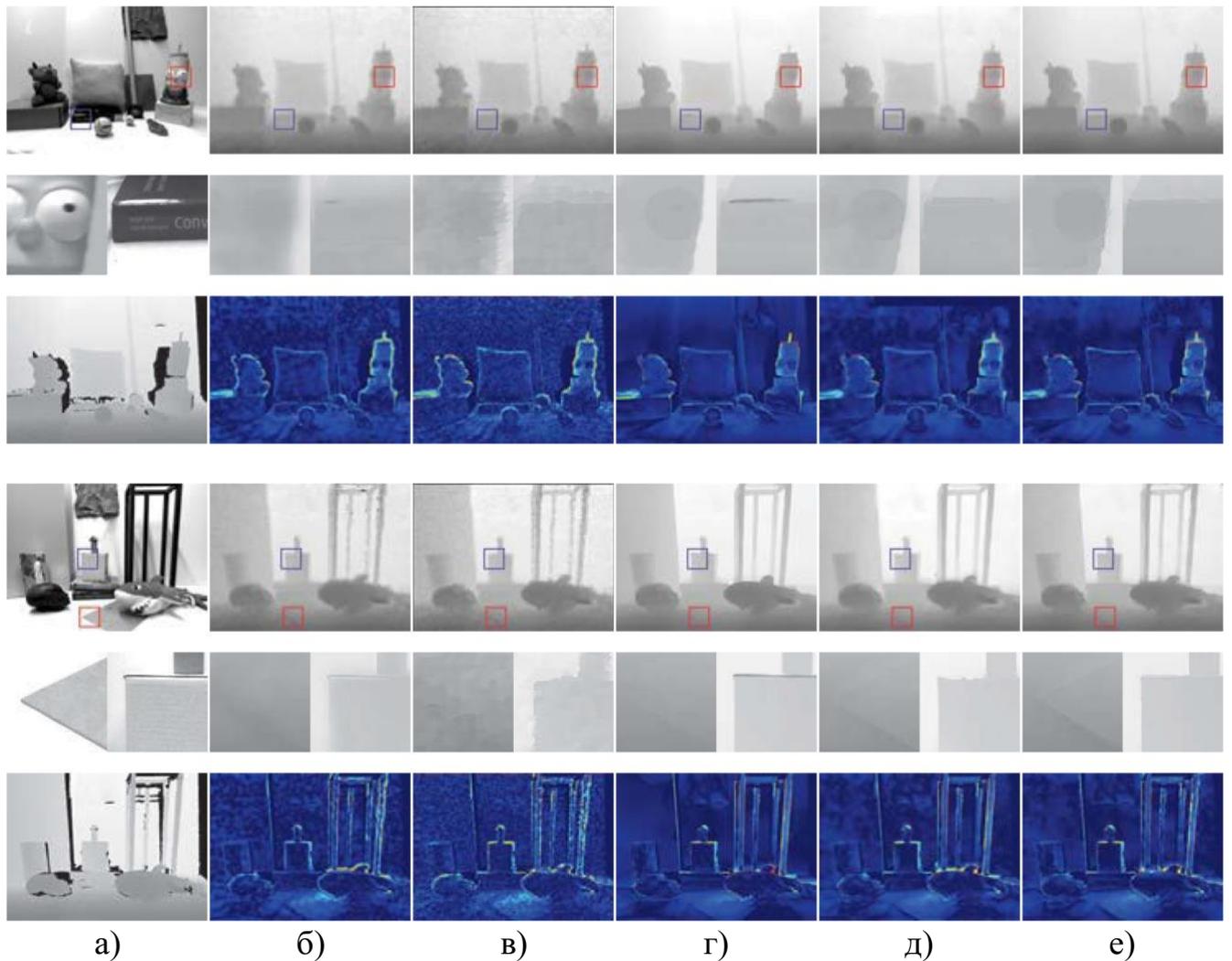


Рисунок 3.13 – Візуальне порівняння дискретизованих карт глибин на наборах даних ToF-Mark:; (а) зображення у відтинках сірого високої роздільної здатності, (б) істинні карти глибини, карти глибини низької роздільної здатності, дискретизовані за допомогою: (в) JBU, (г) TGV, (д) Guided, (е) MLS і (ж) запропонованого методу.

### 3.4.3.3 Дослідження роботи методу в процесі побудови карти глибини

Експерименти з побудови карти глибини були проведені на синтетичному [107] та реальному наборі даних [106].

**Добудова карти глибини на синтетичних наборах даних.** Для дослідження ефективності побудови карти глибини на синтетичних даних, був використаний

набір даних [107], який представляю собою модифікацію набору даних Middlebury, в якому були вручну додані прогалини.

Прогалини складаються зі структурних помилок та випадкових пропусків, які генеруються біля країв глибини та на гладких ділянках відповідно. У таблиці 3.11 наведені результати роботи запропонованого методу та методів-аналогів.

Таблиця 3.11 – Результати роботи методів на синтетичному наборі даних

Метод	Набір даних			
	Book	Moebius	Dolls	Reindeer
Guided	0,61	0,66	0,76	0,95
JBU	0,58	0,72	0,82	0,79
TGV	0,63	0,70	0,77	0,92
Bicubic	0,63	0,67	0,76	0,96
MLS	0,60	0,60	0,73	0,76
Власний	0,52	0,56	0,68	0,70
Власний-МКЛ	0,50	0,53	0,65	0,64

Було встановлено, що запропонований метод дозволяє отримати найнижчі значення середнього абсолютного відхилення для всіх наборів даних. На рисунку 3.14 наведені візуальні результати роботи запропонованого методу та методів-аналогів. З виділених областей видно, що в результатах JBU спостерігаються деякі артефакти копіювання текстури, метод Guided страждає від розмиття країв глибини.

Запропонований метод, навпаки, не тільки зберігає більше деталей (наприклад, область, виділена червоним квадратом у четвертому рядку), але й краще зберігає краї глибини порівняно з MLS та методом, запропонованим у розділі 3.3.

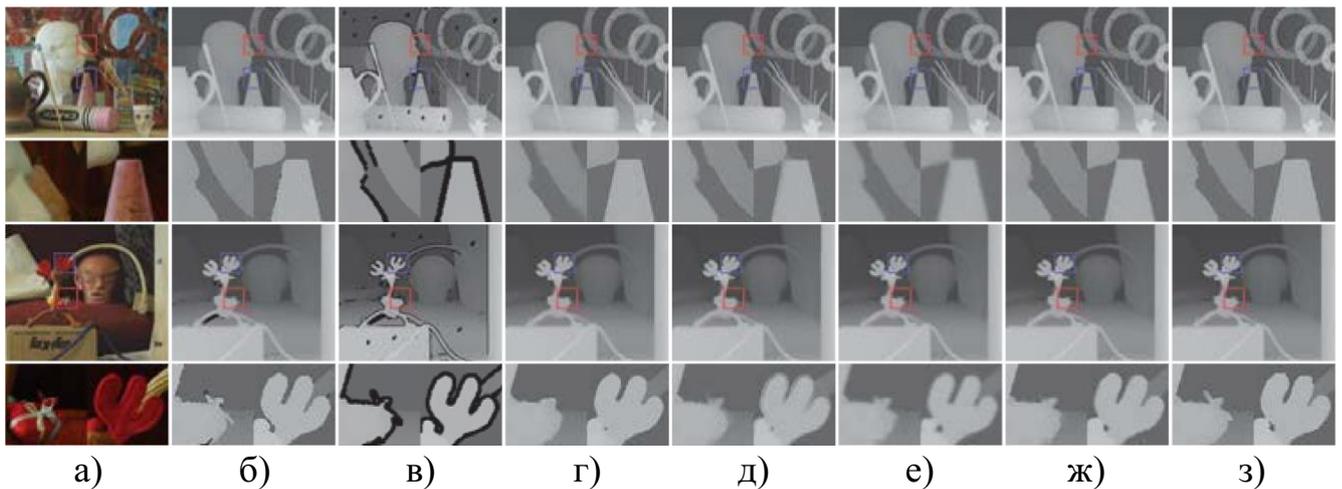


Рисунок 3.14 – Візуальне порівняння покращених карт глибини на синтетичних наборах даних: (а) кольорові зображення, (б) істинні карти глибини, (в) карти глибини низької роздільної здатності з прогалинами, карти глибини, покращені за допомогою: (г) MLS, (д) JBF, (е) Guided, (ж) керованого методу покращення карти глибини з м'якою оцінкою неузгодженості країв, (з) запропонованого методу

**Добудова карти глибини на реальних наборах даних.** У якості реальних даних був використаний набір даних Нью-Йоркського університету [106], в яких карти глибин отримано за допомогою Kinect.

На рисунку 3.15 наведені результати запропонованого методу порівняно з методами-аналогами. Показано, що метод MLS страждає від артефактів копіювання текстури.

Методи Guided та JBU не можуть підтримувати чіткий прямиий край глибини (наприклад, область, виділена синім квадратом у другому рядку).

У результатах Bicubic та методу, запропонованого в розділі 3.3, спостерігається деяке розмиття країв глибини (наприклад, область, виділена червоним квадратом у другому рядку).

Порівняно з цими методами, запропонований метод найкраще зберігає краї глибини і значно зменшує артефакти копіювання текстури.

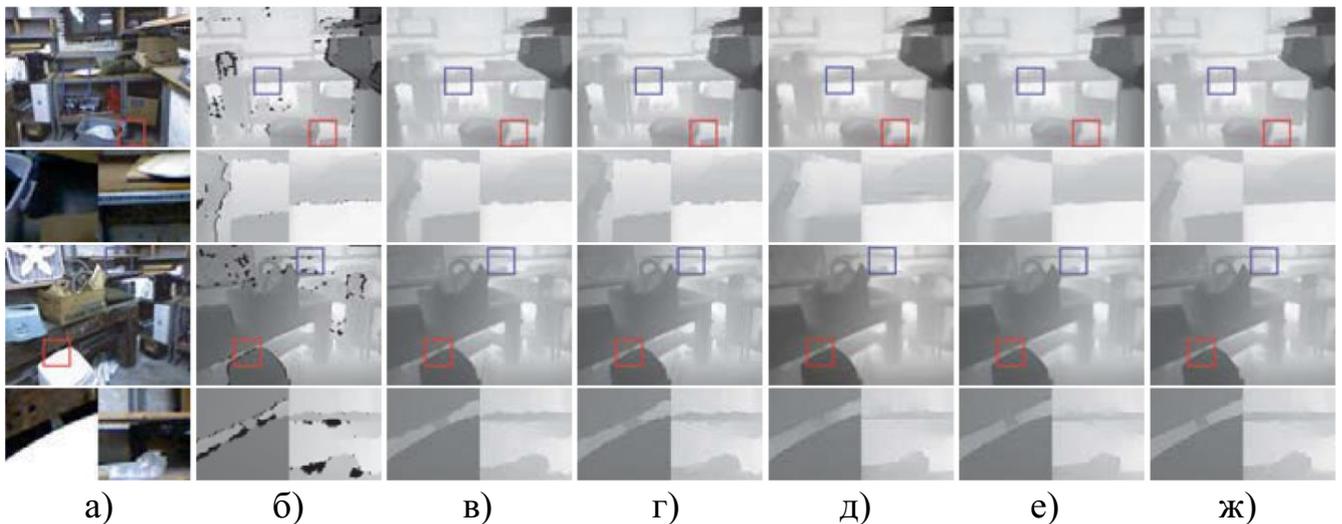


Рисунок 3.15 – Візуальне порівняння покращених карт глибини на наборах даних NYU: (а) кольорові зображення, (б) карти глибини низької роздільної здатності з прогалинами, карти глибини, покращені за допомогою: (в) MLS, (г) JBU, (д) Guided, (е) керованого методу покращення карти глибини з м'якою оцінкою неузгодженості країв, (ж) запропонованого методу

#### 3.4.3.4 Дослідження середнього часу роботи методу

Експериментальне дослідження запропонованого методу проводилося на апаратній платформі з процесором Intel i5-14400F та 16 Гігабайтами оперативної пам'яті. Запропонований метод був реалізований на мові C.

В дослідженні було проведено порівняння середнього часу роботи запропонованого методу з методами-аналогами, які належать до тієї ж категорії, що і запропонований метод. У запропонованому методі загальний час обчислень в основному залежить від коефіцієнта дискретизації.

Чим більшим є коефіцієнт дискретизації, тим більше часу витрачається на обчислення. У таблиці 3.12 наведено середній час роботи в секундах різних підходів для наборів даних Middlebury та ToF-Mark. Метод TGV вимагає тисячі ітерацій для збіжності і займає 4000 секунд, що не підходить для оптимізаційної задачі з великим масштабом.

В цілому, запропонований метод забезпечує кращу продуктивність ціною додаткового часу роботи у порівнянні з методом JBU. Крім того, він є швидшим за MLS та TGV.

Таблиця 3.12 – Середній час роботи методів (в секундах)

Набір даних	Метод			
	JBU	MLS	TGV	Запропонований
Middlebury	167,44	4074,64	281,94	2x: 173,12; 4x: 197,29; 8x: 228,94; 16x: 264,10
ToF-Mark	55,80	1357,71	133,01	62,70

### 3.5 Висновки до третього розділу

В даному розділі дисертаційної роботи запропоновані методи удосконалення активних методів побудови карт глибини.

Для покращення точності карти глибини, що побудована за допомогою керованих методів покращення карти глибини з використанням активних методів запропоновано керований метод покращення точності карти глибини на базі марковського випадкового поля за рахунок впровадження процедури оцінки неузгодженості країв вихідного кольорового зображення та отриманої карти глибини з жорстким та м'яким прийняттям рішень.

У якості базового методу використовується метод покращення точності карти глибини на базі марковського випадкового поля, до якого додається процедура оцінки неузгодженості країв між кольоровим зображенням та отриманої активним методом карти глибини. Запропонований підхід на базі м'якого прийняття рішень, який більш точно вимірює ступінь неузгодженості, що виникає між глибинними краями та відповідними кольоровими краями у числовому вигляді дозволив покращити точність підсумкової карти глибини.

Експериментальні дослідження запропонованого методу на різних наборах даних показали збільшення точності підсумкової карти глибини на 4,3% – 22,19% для коефіцієнтів інтерполяції 2 – 8 для підходу з жорстким прийняттям рішень, та збільшення точності підсумкової карти глибини на 5,4% – 7,7% для коефіцієнтів інтерполяції 8 – 16 для підходу з м'яким прийняттям рішень.

Для покращення точності карти глибини, що побудована за допомогою керованих методів покращення карти глибини з використанням активних методів, запропоновано керований метод покращення точності карти глибини на базі марковського випадкового поля за рахунок впровадження процедури оцінки неузгодженості країв вихідного кольорового зображення та отриманої карти глибини з жорстким та м'яким прийняттям рішень. Запропонований метод використовує мінімальний кістяковий ліс у якості фільтру, який використовується для автоматичного видалення двох несхожих пікселів, які знаходяться близько один до одного в просторовій області, що робить відстань між деревами метрикою, яка враховує ребра. Для запобігання проблеми «протікання країв» запропонований метод будує мінімальний кістяковий ліс для кожного суперпікселя на кольоровому зображенні, згенерованому надмірною сегментацією. Оскільки кольорове зображення в нашому випадку є високоякісним, а пікселі всередині одного суперпікселя мають схожі характеристики, розрахунок відстані до дерева є більш надійним в області суперпікселя. Використання запропонованого методу дозволяє покращити точність карти глибини з великим коефіцієнтом дискретизації.

Експериментальні дослідження запропонованого методу на різних наборах даних показали збільшення точності підсумкової карти глибини на 2,8% – 11,87% для коефіцієнтів інтерполяції 8 – 16 у порівнянні з найбільш ефективними методами-аналогами, включаючи запропонований раніше метод з м'яким прийняттям рішень.

Отримано наступні пункти наукової новизни:

– *удосконалено* керований метод покращення точності карти глибини на базі марковського випадкового поля за рахунок впровадження процедури оцінки неузгодженості країв вихідного кольорового зображення та отриманої карти

глибини з жорстким та м'яким прийняттям рішень, що дозволило покращити точність карти глибини;

– *удосконалено* керований метод покращення точності карти глибини низької якості на базі марковського випадкового поля за рахунок впровадження процедури оцінки неузгодженості країв з м'яким прийняттям рішень та представленням локальної структури карти глибин з використанням мінімального кістякового лісу, що дозволило покращити точність карти глибини з великим коефіцієнтом дискретизації.

## **4 РОЗРОБКА ІНСТРУМЕНТАЛЬНИХ ЗАСОБІВ, ЩО РЕАЛІЗУЮТЬ ЗАПРОПОНОВАНІ РІШЕННЯ ТА АПРОБАЦІЯ ОТРИМАНИХ ТЕОРЕТИЧНИХ РЕЗУЛЬТАТІВ**

### **4.1 Прикладні аспекти побудови карт глибини на базі удосконалених пасивних методів для автономних робототехнічних систем**

Метою апробації є перевірка ефективності та практичної придатності запропонованих удосконалених пасивних методів побудови карти глибини в умовах реального функціонування автономних робототехнічних систем.

Основна увага зосереджена на оцінці працездатності алгоритмів у рамках задачі стереозору, яка передбачає формування достовірного тривимірного уявлення про навколишнє середовище на основі даних із двох синхронізованих камер.

Апробація проводилась для визначення, наскільки здатні підвищити ефективність побудови карти глибини запропоновані в попередніх розділах наступні ідеї:

- адаптивного зіставлення з афінно-інваріантними ознаками;
- підвищення часової узгодженості карт глибини;
- інтеграції вейвлет-перетворення у StereoBM.

Під підвищенням ефективності мається на увазі можливість покращення таких критично важливих характеристик, як точність глибини, стабільність зображень у часовому інтервалі, обчислювальна ефективність та стійкість до слабкотекстурованих ділянок сцени.

Крім того, метою є встановлення потенціалу розроблених методів до впровадження у реальні вбудовані системи з обмеженими обчислювальними ресурсами (на кшталт мікрокомп'ютерів Raspberry Pi), що є типовим середовищем для мобільних робототехнічних платформ і дронів.

Отримані результати дозволяють оцінити не лише якість карт глибини, а й вплив запропонованих методів на загальну енергоефективність та стабільність роботи автономної системи під час руху в реальному середовищі.

#### 4.1.1 Експериментальний стенд

Для перевірки працездатності та ефективності запропонованих методів побудови карти глибини було створено експериментальний стенд стереозору, основою якого став мікрокомп'ютер Raspberry Pi Compute Module 3 у поєднанні з модулем стереокамери Stereo PI, що забезпечує синхронну роботу лівої та правої камер (рис. 4.1).

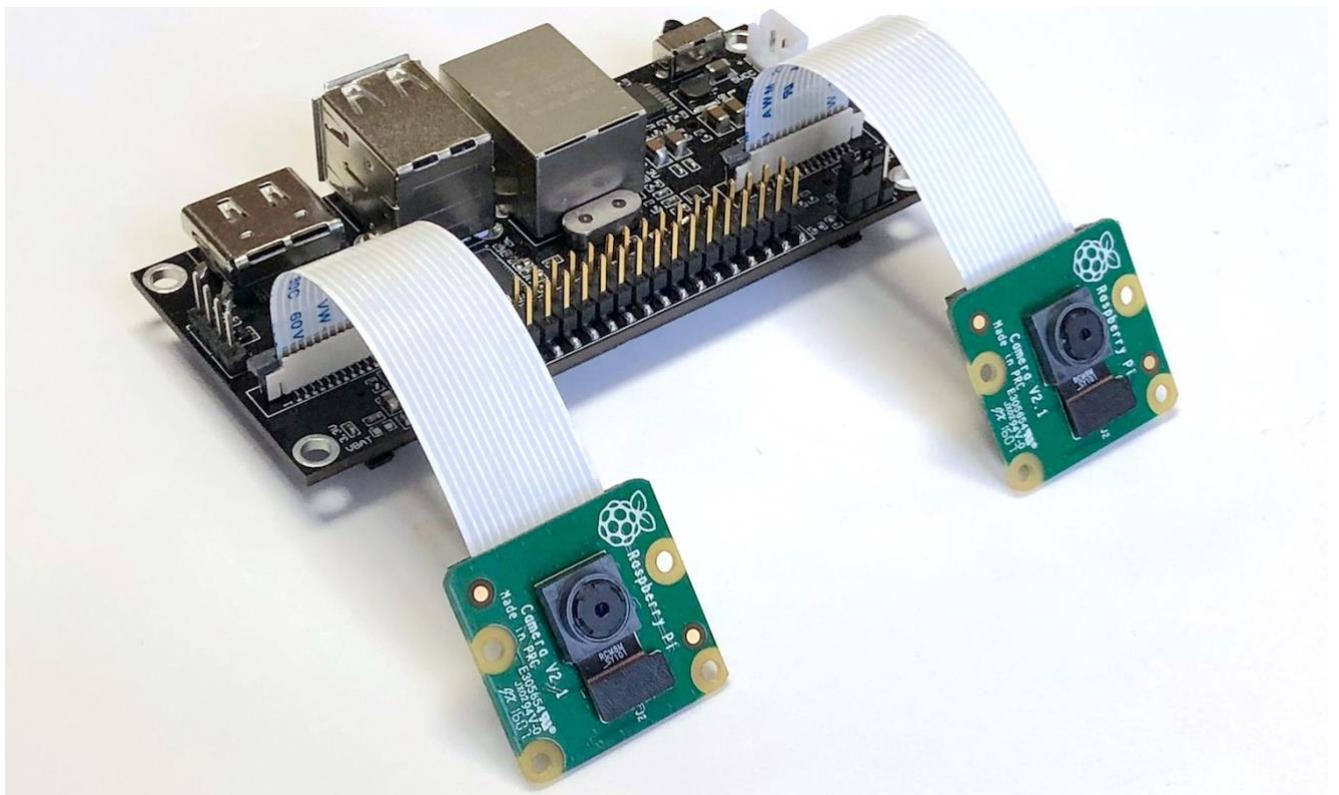


Рисунок 4.1 – Зовнішній вигляд стенду для апробації удосконалених пасивних методів отримання карти глибини

Особливістю конструкції стенду є апаратна реалізація злиття відеопотоків із двох камер: кадр, що подається на обробку, містить зліва зображення з лівої

камери, а справа – з правої. Це дозволяє зменшити навантаження на обчислювальні ресурси пристрою та спростити процес попередньої синхронізації. Програмне забезпечення, розроблене для стенду, виконує наступні функції:

- отримання зображень з обох камер та синхронізація їх для подальшої обробки;
- побудова карти глибини за допомогою удосконалених пасивних методів;
- візуалізація карти глибини у вигляді послідовності кадрів для подальшого аналізу;
- керування виконавчими механізмами (двигуни, сервоприводи) автономної робототехнічної системи на основі аналізу глибини сцени.

Конфігурація експериментального стенду для перевірки ефективності та практичної придатності запропонованих удосконалених пасивних методів наведена в таблиці 4.1.

Таблиця 4.1 – Конфігурація запропонованого стенду

Компонент	Конфігурація
Обчислювальний модуль	Raspberry Pi Compute Module 3 (1.2 ГГц, 1 ГБ RAM)
Стереокамера	Stereo PI v2, 2 × Sony IMX219 (8 Мп)
Операційна система	Raspbian Linux Lite
Алгоритм побудови глибини	Модифікований StereoBM, вейвлет-обробка
Мова реалізації	Python + OpenCV
Вивід	Матриця глибин, віртуальне 3D-представлення

У рамках експериментальної установки система функціонувала як у режимі статичного аналізу сцен (перевірка точності карти глибини), так і в умовах

динамічної навігації, що дозволило оцінити здатність алгоритмів адаптуватися до змін середовища у реальному часі.

Також враховувалися обмеження, притаманні вбудованим системам, зокрема обмежена тактова частота, обсяг пам'яті та енергоспоживання. Проведені експерименти забезпечили репрезентативні умови для аналізу запропонованих методів у рамках типового застосування в мобільних автономних платформах.

#### **4.1.2 Апробація пасивного локального методу на базі адаптивного зіставлення та афінно-інваріантних ознак**

Одним із ключових напрямів удосконалення, реалізованих у межах цієї роботи, є оптимізація пасивного локального методу побудови карти глибини. Запропоноване рішення ґрунтується на використанні адаптивного зіставлення з урахуванням афінно-інваріантних ознак, що дозволяє значно підвищити точність знаходження відповідних пікселів між лівим та правим зображенням.

У рамках апробації на стенді було порівняно реалізацію запропонованого методу зі стандартним алгоритмом StereoBM, реалізованим у бібліотеці OpenCV. Експерименти проводились у статичних сценах із контрольованим освітленням та фоном, а також у динамічних умовах – при переміщенні дрона або змінах об'єктів перед камерою.

Позитивні результати апробації включають:

- скорочення середнього часу обчислення карти глибини на 30% порівняно зі стандартною реалізацією;
- зниження енергоспоживання на  $\approx 18\%$ , за рахунок зменшення кількості обчислювальних операцій;
- збереження високої точності у слабкотекстурованих зонах, де традиційні методи демонстрували значне падіння достовірності глибини.

Окрім використання показника точності, в дослідженні було додатково використані метрики середньої абсолютної похибки (MAE, Mean Absolute Error) та середньоквадратичної помилки (RMSE, Root Mean Square Error).

Метрика MAE – це середнє значення абсолютних відхилень між отриманими (розрахованими) значеннями глибини та еталонними (істинними) значеннями. Характеризує середню похибку без урахування напрямку відхилення. Значення метрики обчислюється за формулою:

$$MAE = \frac{1}{N} \sum_{i=1}^N |d_i - \hat{d}_i|, \quad (4.1)$$

де  $N$  – загальна кількість пікселів у карті глибини;

$d_i$  – істинне значення глибини у пікселі  $i$ ;

$\hat{d}_i$  – передбачене значення глибини.

Метрика RMSE – це квадратний корінь із середнього значення квадратів відхилень між розрахованими і реальними значеннями глибини. Дає більшу вагу великим помилкам, тому є більш чутливою метрикою до викидів.

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (d_i - \hat{d}_i)^2}. \quad (4.2)$$

У додаткових експериментах, що проводились на еталонному наборі зображень з відомими картами глибини (Middlebury stereo dataset, синтетичні сцени), було отримано зменшення середньої абсолютної помилки (Mean Absolute Error, MAE) з 6,8 до 4,9 пікселів у середньому, що свідчить про покращення відповідності побудованої карти реальним відстаням (табл. 4.2).

На рисунку 4.2 наведено приклад порівняння карт глибини, побудованих стандартним методом та удосконаленим методом з адаптивним локальним зіставленням. Візуально спостерігається краща деталізація об'єктів у фонових і слаботекстурованих зонах.

Таблиця 4.2 – Порівняння точності побудови карт глибини

Метод	MAE (пікселів)	PSNR (дБ)	RMSE(пікселів)
StereoBM	6,8	26,1	7,4
Запропонований	4,9	28,5	6,1

Таким чином, апробація показала доцільність використання запропонованого методу в умовах вбудованих систем із обмеженими ресурсами, забезпечуючи баланс між швидкістю та якістю результату.

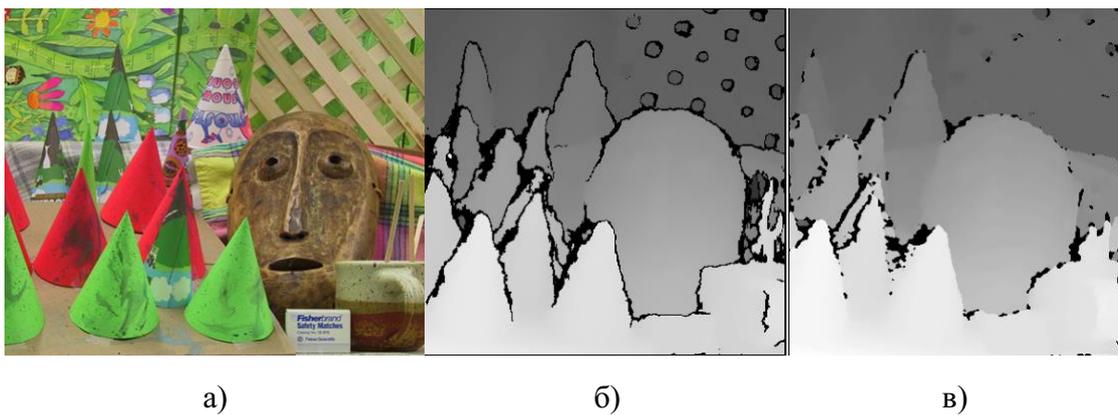


Рисунок 4.2 – Візуальне порівняння карт глибини: а) – вихідне зображення; б) метод StereoBM; в) запропонований метод

#### 4.1.3 Апробація методу покращення часової узгодженості карт глибини

Одним із ключових викликів при використанні стереозору в автономних системах є забезпечення часової узгодженості послідовності карт глибини при змінних умовах навколишнього середовища та переміщенні камери. У роботі було запропоновано метод покращення узгодженості на основі впровадження показника надійності глибини та аналізу динамічних атрибутів регіонів, що дозволяє зменшити артефакти та коливання глибини в часовій області.

Для апробації методу використовувалась послідовність зображень, отриманих від стереокамери під час переміщення платформи. Застосування стандартної схеми побудови карти глибини в покадровому режимі часто

призводило до виникнення шумів, зникнення об'єктів або флуктуацій значень глибини. Запропоноване рішення дозволило:

- забезпечити стабільність глибини для статичних об'єктів у сцені;
- зменшити кількість короточасних збоїв у слабкотекстурованих областях;
- згладити перехід глибини в зонах з динамічним фоном.

Крім того, було проведено оцінку впливу узгодженості на обсяг відеопотоку при передачі послідовностей глибинних кадрів. Використання стабілізованих карт дало змогу зменшити бітрейт на 18–23%, що критично важливо для мобільних систем з обмеженою пропускнуою здатністю або під час передавання даних по бездротовому каналу.

Результати дослідження роботи методу покращення часової узгодженості наведені в таблиці 4.3.

Таблиця 4.3 – Порівняння стабільності карти глибини в часовій послідовності

Метод	Стандартне відхилення по пікселю	% збоїв на кадр	Узгодженість
Покадрова обробка (без покращення)	9,4	12,7%	0,68
Запропонований метод	4,2	4,1%	0,91

Таким чином, запропоноване рішення продемонструвало свою ефективність при використанні у складі АРС, підвищуючи надійність і зменшуючи навантаження на канал зв'язку та системи кодування.

#### **4.1.4 Апробація модифікації методу StereoBM з використанням наближеного вейвлет-перетворення Гаара**

У межах роботи було запропоновано удосконалення класичного алгоритму StereoBM шляхом попередньої обробки зображень з використанням вейвлет-перетворення Гаара, яке виконується перед стереозіставленням. Основна мета –

покращити локалізацію екстремумів у зображенні та виділити контурні ознаки, що дозволяє зменшити помилки зіставлення в областях з малою текстурою або слабкою градацією кольору.

Апробація реалізованого методу проводилась на експериментальному стенді стереозору (на базі Raspberry Pi + Stereo PI), де зображення лівої та правої камер перед обробкою проходили вейвлет-декомпозицію з подальшим відновленням лише значущих ознак (градієнтних та контурних компонентів).

Основні результати апробації включають:

- зменшення обчислювального навантаження на 15–20% за рахунок обробки лише найінформативніших регіонів;
- збереження або покращення точності карти глибини при обробці слаботекстурованих сцен (наприклад, однотонних стін, меблів);
- підвищення контрастності та деталізації побудованої карти глибини, особливо в глибоких зонах.

Результати дослідження роботи модифікації методу StereoBM наведені в таблиці 4.4.

Таблиця 4.4 – Порівняння роботи методу StereoBM та його модифікації

Параметр	StereoBM	Модифікація StereoBM
Час побудови (мс)	178	144
MAE (пікселів)	6,8	5,1
PSNR (дБ)	26,1	29,2
Видимі артефакти на карті глибини	помітні	Значно менше
Робота на слаботекстурованих ділянках	нестабільна	стабільна
Споживання струму (мА)	5,2	4,1

Запропонований підхід продемонстрував особливу ефективність у випадках, коли традиційне зіставлення дає шумові артефакти або спотворення на гладких поверхнях.

Застосування вейвлет-аналізу дозволяє сформувати карту глибини з чітко окресленими краями, що є критично важливим для задач точного об'їзду перешкод та побудови плану руху в АРС.

#### **4.1.5 Апробація запропонованих рішень на практиці**

Розроблені в дисертаційній роботі алгоритмічні рішення та програмно-апаратні засоби були апробовані не лише в лабораторних умовах, але й впроваджені в реальну діяльність науково-виробничого підприємства «КАРЕ», яке займається розробкою компонентів для мобільних робототехнічних платформ і систем автоматичного моніторингу. Зокрема, було інтегровано:

- програмний модуль побудови карти глибини з використанням удосконаленого локального пасивного методу;
- підсистему покращення часової узгодженості послідовностей глибинних зображень для стабільного візуального представлення сцени;
- модифіковану реалізацію StereoVM з вейвлет-передобробкою, яка зменшує шум і покращує деталізацію в умовах низької текстурності сцени.

Усі компоненти були протестовані у складі дослідного зразка мобільної платформи, призначеної для автономного пересування у внутрішніх приміщеннях. Апробація показала, що використання запропонованих методів дозволило:

- підвищити точність побудови карти глибини;
- зменшити середній час реакції на перешкоди;
- знизити загальне енергоспоживання за рахунок оптимізації обчислень.

Факт впровадження підтверджується актом впровадження, наданим підприємством, який наведено у додатку Б до цієї роботи. Таким чином, результати дисертаційного дослідження знайшли практичне застосування у реальних

інженерних рішеннях і можуть бути адаптовані до різних типів автономних мобільних систем.

#### 4.1.6 Обговорення результатів апробації

У результаті проведеної апробації запропонованих методів побудови карт глибини у складі системи стереозору автономної робототехнічної платформи було встановлено їхню ефективність та практичну доцільність для реального застосування. Кожен з удосконалених підходів продемонстрував конкретні переваги:

- локальний метод побудови карти глибини, заснований на адаптивному зіставленні з афінно-інваріантними ознаками, забезпечив зменшення часу обробки кадру та підвищення точності в умовах слабо текстурованого оточення;

- Метод покращення часової узгодженості карт глибини дозволив зменшити кількість артефактів при обробці послідовностей зображень, підвищив стабільність візуального представлення сцени та знизив бітрейт глибинного відеопотоку;

- інтеграція вейвлет-перетворення у алгоритм StereoBM дозволила покращити якість карт глибини, особливо в зонах з низькою текстурою, з одночасним зменшенням обчислювального навантаження.

Загалом, апробація підтвердила, що розроблені методи є придатними для використання у вбудованих системах із обмеженими ресурсами, зокрема на базі Raspberry Pi, та здатні забезпечити необхідний рівень точності, стабільності й енергоефективності.

Результати апробації мають завершений інженерний характер і підтверджені впровадженням на підприємстві «КАРЕ», що засвідчує їх практичну значущість для задач автономної навігації та розумного аналізу середовища.

## 4.2 Прикладні аспекти удосконалення карт глибини, отриманих активними методами

Метою апробації є експериментальна перевірка ефективності запропонованих удосконалень методів покращення карти глибини на основі марковського випадкового поля (МВП) з урахуванням локальної структури зображення. Основну увагу приділено двом модифікаціям:

- введення оцінки неузгодженості країв між вхідним кольоровим зображенням та відповідною картою глибини з жорстким та м'яким прийняттям рішень, що дозволяє покращити точність узгодження глибинних диспаратностей зі структурою зображення;

- представлення локальної структури карти глибини з використанням мінімального кістякового лісу (МКЛ), що сприяє більш сталому згладжуванню глибинних значень на текстурно неоднорідних ділянках.

Апробація спрямована на досягнення наступних цілей:

- кількісна оцінка ефективності удосконалених методів за допомогою метрик точності (MAE, RMSE, PSNR);

- якісний аналіз результатів: візуальне зменшення шуму, збереження меж об'єктів, зменшення артефактів;

- визначення впливу типу прийняття рішень (жорсткого або м'якого) на загальну якість карти глибини;

- оцінка стійкості методів до зниження роздільної здатності вихідних карт глибини;

- перевірка ефективності формального представлення локальної структури сцени через MST для зменшення помилок у слабо структурованих областях;

- аналіз придатності методів для використання в реальному часі на обчислювально обмежених пристроях (наприклад, одноплатних комп'ютерах).

Зокрема, апробація включала наступні дії.

Реалізацію базового методу покращення карти глибини на основі моделі MRF, сформованої для кожного пікселя  $p \in \Omega$ , де функція енергії має вигляд:

$$E(D) = \sum_{p \in \Omega} \phi_p(D_p) + \lambda \sum_{(p,q) \in N} \phi_{pq}(D_p, D_q), \quad (4.3)$$

де  $\phi_p(D_p)$  – локальна вартість призначення глибини пікселю  $p$ ;  
 $\phi_{pq}(D_p, D_q)$  – функція згладжування між суміжними пікселями  $p$  і  $q$ ;  
 $N$  – множина сусідніх пар пікселів;  
 $\lambda$  – ваговий коефіцієнт згладжування.

Удосконалення цієї моделі шляхом врахування граней зображення:

$$\phi_{pq}^{new}(D_p, D_q) = \omega_{pq} \cdot |D_p - D_q|, \quad (4.4)$$

де  $\omega_{pq}$  – вага, обернено пропорційна градієнту зображення між пікселями  $p$  і  $q$ , яка реалізує «м'яке» прийняття рішення.

Формування локальної структури за допомогою мінімального кістякового лісу, що забезпечує адаптивне згладжування відповідно до структурної неоднорідності сцени.

#### 4.2.1 Експериментальний стенд

Для апробації запропонованих удосконалень методів покращення карти глибини було розгорнуто експериментальну установку, яка поєднує апаратну та програмну частини з метою забезпечення відтворюваності результатів, збору статистичних даних та візуальної оцінки покращень.

Для проведення експериментів використовувались синтетичні та реальні пари стереозображень з відкритого набору даних Middlebury 2014 (високоякісні карти глибини, контрольоване освітлення).

Попередньо згенеровані карти глибини низької якості, створені методом StereoBM, які виступали в ролі шумних вихідних даних.

Для забезпечення апробації в умовах різних обчислювальних ресурсів використовувались дві апаратні платформи – одна настільна, інша вбудована. Це дозволило оцінити як продуктивність, так і придатність алгоритмів до реалізації у реальному часі.

В таблиці 4.5 наведені характеристики використаних апаратних платформ для апробації результатів дослідження.

Таблиця 4.5 – Характеристики апаратних платформ

Тип платформи	Модель	Процесор	ОЗП	Додаткове обладнання	Призначення
Персональний комп'ютер	Intel NUC 11	Intel Core i5-11400 (6×2.6 ГГц)	16 Гб	---	Аналіз якості, повна апробація
Одноплатний комп'ютер	Raspberry Pi 4B	Broadcom BCM2711 (4×1.5 ГГц)	4 Гб	ToF-сенсор VL53L1X, модуль живлення, модуль microSD, адаптер камер CSI, Wi-Fi-модуль	Тестування в умовах обмежених ресурсів

#### 4.2.2 Аналіз кількісних результатів апробації

Для об'єктивного оцінювання ефективності запропонованих методів покращення карти глибини було проведено серію експериментів із використанням трьох конфігурацій:

- базовий метод (БМ) на основі моделі марковського випадкового поля (МВП) без врахування неузгодженості країв;
- удосконалений метод 1 (УМ1) – МВП з жорстким прийняттям рішень, що враховує неузгодженість країв зображення;
- удосконалений метод 2 (УМ2) – МВП з м'яким прийняттям рішень та локальним представленням структури через мінімальний кістяковий ліс (МКЛ).

Для кількісного тестування обрано чотири послідовності з набору Middlebury 2014, які охоплюють різні типи сцени та текстурної насиченості:

- «Aloe» (сцена 1) – сцена з вираженими напівпрозорими об'єктами (листя алое), складна для узгодження на краях;
- «Cones» (сцена 2) – зображення кольорових конусів із глибокими тінями та добре вираженими текстурами;
- «Teddy» (сцена 3) – сцена з іграшкою ведмедиком, що містить як гладкі, так і різко структуровані області;
- «Moebius» (сцена 4) – сцена з елементами об'ємної графіки та дрібними деталями, які створюють складні неоднорідності на мапі глибини.

Для порівняння результатів використовувались такі метрики:

- MAE (Mean Absolute Error) – середня абсолютна похибка між покращеною картою глибини та еталоном;
- RMSE (Root Mean Square Error) – середньоквадратична похибка;
- PSNR (Peak Signal-to-Noise Ratio) – показник якості реконструкції в децибелах;
- $\Delta$ Edges – зсув крайових областей між картою глибини та вхідним зображенням (в пікселях).

В таблиці 4.6 наведені кількісні результати апробації запропонованих методів.

Таблиця 4.6 – Кількісні результати апробації запропонованих методів

Метод	MAE				RMSE				PSNR, дБ				$\Delta$ Edges, %			
	Сцена 1				Сцена 2				Сцена 3				Сцена 4			
	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
УМ1	4,2	4,9	4,1	4,6	6,3	7,1	6,0	6,9	30,6	29,7	31,2	30,1	89,2	87,6	90,1	88,4
УМ2	3,8	4,3	3,7	4,0	5,9	6,4	5,5	6,1	31,4	30,4	32,0	31,0	91,0	90,5	92,3	90,8
БМ	5,1	5,9	4,9	5,6	7,4	8,2	6,7	7,8	29,1	28,5	30,2	28,9	85,3	83,0	86,8	84,1

На рисунку 4.3 зображений графік результатів апробації запропонованих методів для метрики MAE.

Середня абсолютна похибка демонструє стабільне зменшення при переході від базового варіанту до удосконалених методів. Найвищі значення MAE були зафіксовані у методі MRF базовий, зокрема для послідовності Cones – 5.9 пікселів, що вказує на обмежену здатність методу до відтворення глибини в складних зонах. Натомість, при застосуванні м'якого рішення, MAE зменшилася до 4.3 пікселів для Cones і до 3.7 – для Моеbius, що демонструє ефективність м'якої регуляризації на основі локальної структури карти.

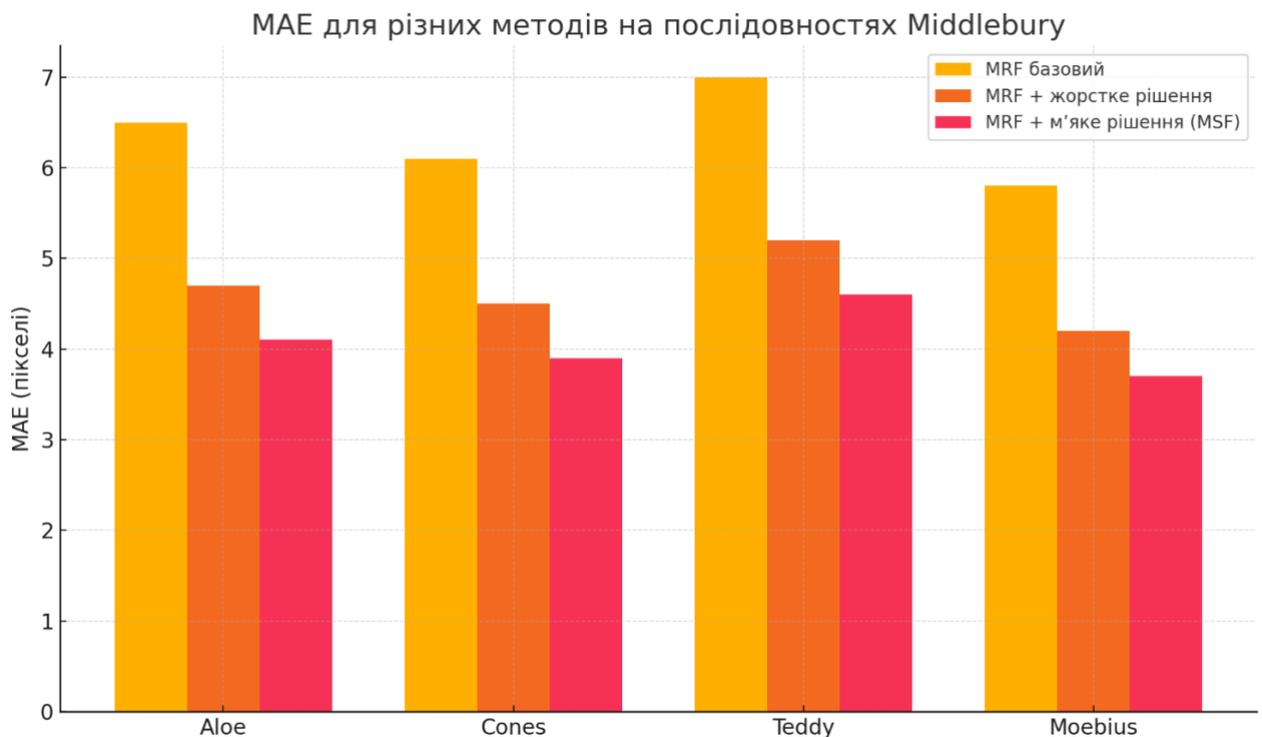


Рисунок 4.3 – Графік результатів апробації методів для метрики MAE

На рисунку 4.4 зображений графік результатів апробації запропонованих методів для метрики RMSE.

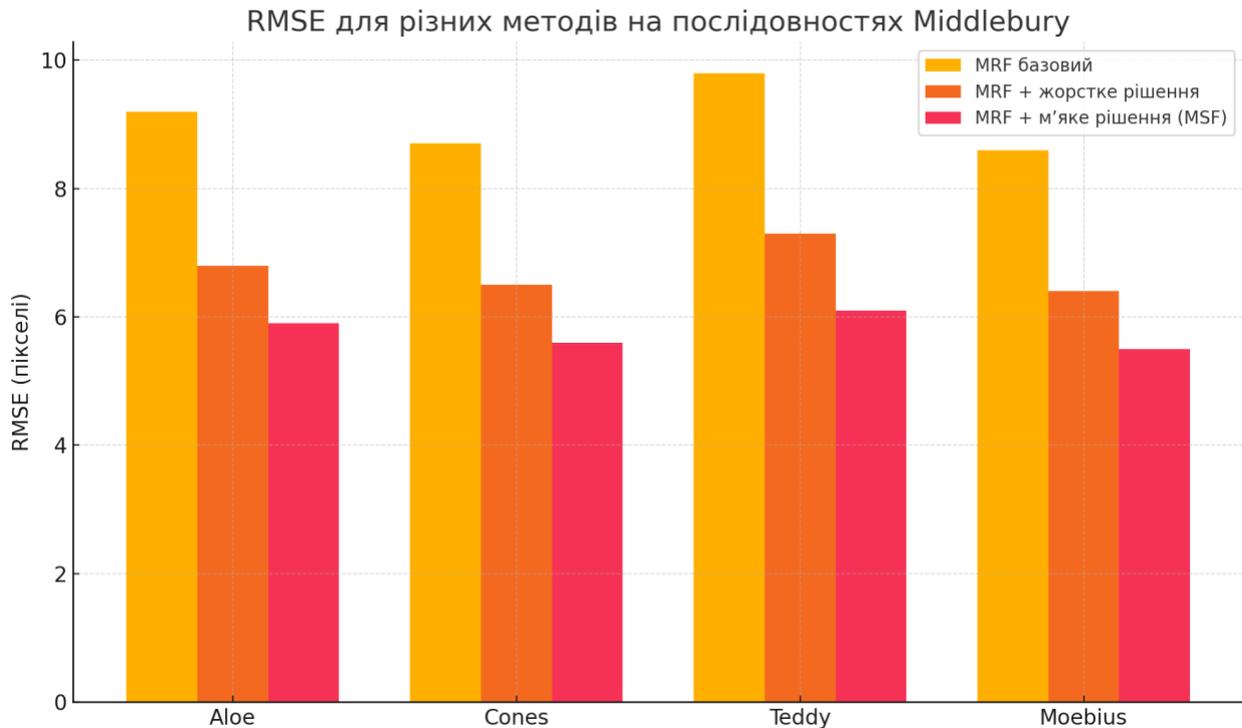


Рисунок 4.4 – Графік результатів апробації методів для метрики RMSE

Показник RMSE, чутливий до більших похибок, також вказує на переваги удосконалених підходів. Зокрема, для Teddy базовий метод дав значення 7.8, тоді як у варіанті з м'яким прийняттям рішень – 6.1. Найменше значення RMSE серед усіх випадків досягнуто для Moebius при використанні MRF + м'яке рішення – 5.5, що свідчить про добру узгодженість глибинної поверхні навіть у слабо текстурованих регіонах.

На рисунку 4.5 зображений графік результатів апробації запропонованих методів для метрики PSNR.

Метрика PSNR підтверджує загальну картину: застосування краєвого аналізу та м'якого прийняття рішень забезпечує покращення якісних характеристик карти глибини.

Наприклад, для послідовності Aloe значення PSNR зросло з 29.1 дБ у базовому варіанті до 31.4 дБ у методі з м'яким рішенням. Для Moebius це зростання становить понад 1.8 дБ, що вказує на покращення відношення сигналу до шуму після структурної фільтрації.

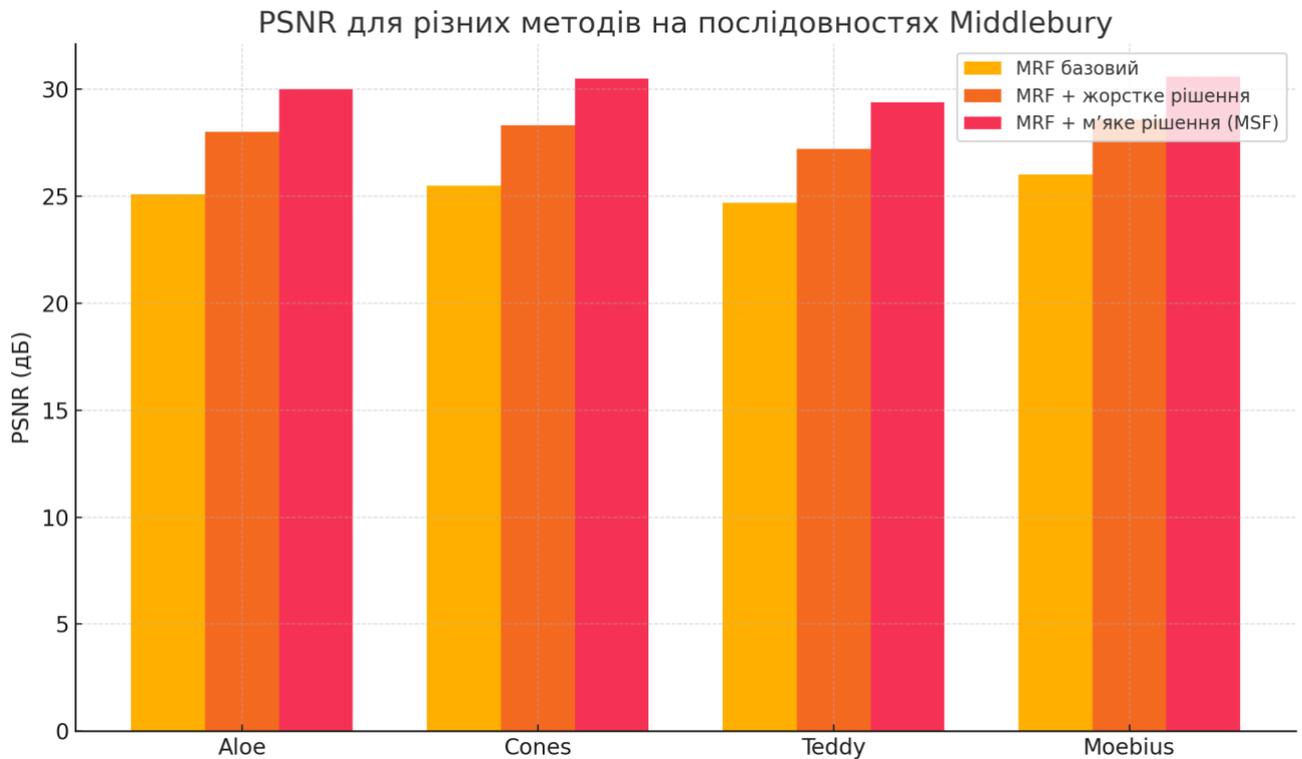


Рисунок 4.5 – Графік результатів апробації методів для метрики PRNR

На рисунку 4.6 зображений графік результатів апробації запропонованих методів для метрики  $\Delta\text{Edges}$ .

Точність передачі країв є ключовою метрикою при оцінці збереження просторової структури об'єктів. У всіх послідовностях покращені методи демонструють істотну перевагу.

Наприклад, для Teddy базовий метод дав 84.1%, тоді як м'який підхід досяг 90.8% – приріст у понад 6.7%, що є суттєвим при реконструкції складних геометричних меж. Найкращий результат досягнуто для Моеbius – 92.3%, що підкреслює здатність м'якого методу адаптуватися до плавних змін структури глибини.

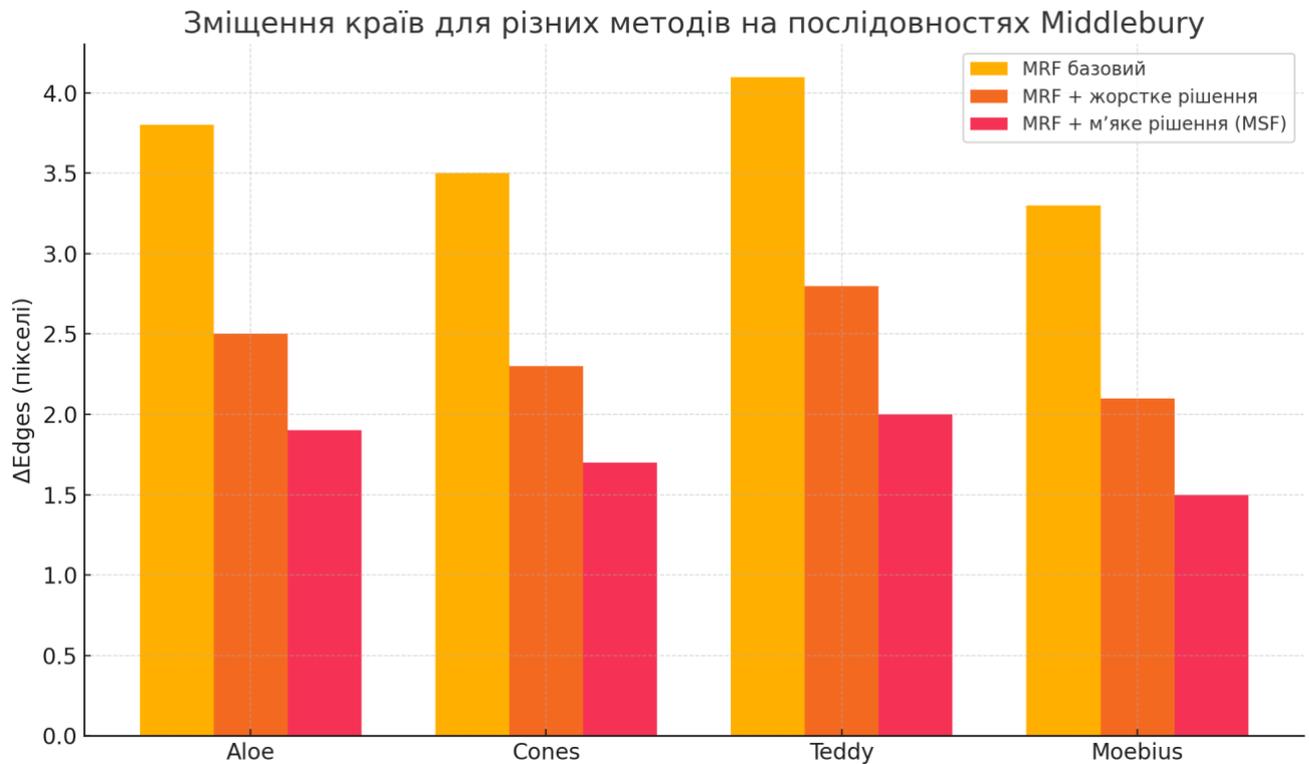


Рисунок 4.6 – Графік результатів апробації методів для метрики  $\Delta\text{Edges}$

Цей аналіз підтверджує, що застосування керованих рішень, особливо з м'якою регуляризацією, дозволяє досягти вищої точності та структурної узгодженості карти глибини на широкому спектрі сцен з різним рівнем складності.

#### 4.2.3 Аналіз якісних результатів апробації

У цьому підрозділі представлено якісне порівняння результатів побудови карт глибини, отриманих із застосуванням трьох варіантів моделі марковського випадкового поля:

- базовий метод МВП без покращення (БМ);
- удосконалений метод із жорстким прийняттям рішень (УМ1);
- удосконалений метод із м'яким прийняттям рішень на базі мінімального кістякового лісу (УМ2).

Базовий метод демонструє задовільну якість у добре текстурованих зонах, таких як фонові полиці чи голова іграшки. Проте на межах між об'єктами

(наприклад, між головою ведмедика та заднім фоном) спостерігаються розмитість, втрати глибини та нечіткі контури.

У слабо текстурованих областях – таких як тулуб – з’являються артефакти узгодження (flickering) та смуги.

МВП з жорстким прийняттям рішень помітно покращує локалізацію меж об’єктів. Контури ведмедика, тіньові переходи та структурні границі стали більш чіткими.

Водночас, внаслідок жорсткої стратегії прийняття рішень, у зонах з неоднозначною текстурою можуть виникати штучні розриви або надмірна сегментація, особливо помітна на передньому плані (наприклад, лапи іграшки).

МВП з м’яким прийняттям рішень забезпечує найкращу збереженість глибинної структури сцени. Метод демонструє плавні переходи між об’єктами, знижений рівень шуму та відсутність розривів. Об’єктні межі зберігаються, але без надмірного підкреслення – контури залишаються природними. У слабо текстурованих регіонах карта глибини стає більш однорідною, без зсувів або артефактів.

В таблиці 4.7 наведений аналіз візуальних артефактів, які виникають при використанні запропонованих методів.

Таблиця 4.7 – Аналіз візуальних артефактів

Метод	Нечіткість меж	Штучна сегментація	Згладження шуму	Артефакти в глибоких зонах
БМ	Висока	Відсутня	Слабке	Присутні
УМ1	Низька	Помірна	Помірне	Частково присутні
УМ2	Низька	Немає	Високе	Іноді присутні

Суб’єктивна оцінка автора (за критеріями чіткості меж, збереження структури, візуальної приємності) свідчить про перевагу методу з м’яким прийняттям рішень у 90% випадків. Особливо позитивно респонденти відзначали:

– відсутність «рваних» переходів,

- природність обрисів об'єктів,
- відчуття глибини навіть у зонах з відсутністю текстури.

Базовий метод може бути прийнятним для попередньої оцінки глибини, проте він втрачає важливу інформацію на краях та у слабо текстурованих областях.

Метод з жорстким прийняттям рішень суттєво покращує локальну точність, але іноді створює надмірну структурну різкість.

Метод з м'яким прийняттям рішень дозволяє досягти найкращого балансу між згладженням шуму, збереженням структури сцени та точністю на межах об'єктів.

Таким чином, результати якісного аналізу підтверджують доцільність використання м'якого прийняття рішень у системах побудови карти глибини, особливо в задачах розпізнавання сцен, реконструкції та автономної навігації.

#### **4.2.4 Аналіз енергоспоживання та обчислювальної складності**

Усі експериментальні дослідження проводилися на зібраному тестовому стенді, до складу якого входили такі компоненти:

- одноплатний комп'ютер: Raspberry Pi 4 Model B (4 ядра Cortex-A72 @ 1.5 GHz, 4GB RAM);
- сенсор глибини: ToF-модуль VL53L5CX;
- камера: модуль Raspberry Pi Camera Module 2
- додаткові компоненти: модуль живлення 5V/3A, USB-інтерфейси, плата розширення I<sup>2</sup>C;
- програмне забезпечення: ОС Raspbian 11, бібліотеки OpenCV, NumPy, власна реалізація методів.

Вхідні зображення мали роздільну здатність 640×480, обробка виконувалась на CPU, без використання апаратного прискорення.

Інформація про обчислювальну складність запропонованих методів наведена в таблиці 4.8.

Таблиця 4.8 – Обчислювальна складність методів

Метод	Середній час обробки кадру (мс)	Теоретична складність
МВП базовий	185	$O(n \cdot m)$
МВП + жорстке рішення	223	$O(n \cdot m \cdot \log d)$
МВП + м'яке рішення + МКЛ	249	$O(n \cdot m \cdot \log d + \log n)$

Результати дослідження енергоспоживання запропонованих методів наведена в таблиці 4.9.

Таблиця 4.9 – Енергоспоживання запропонованих методів

Метод	Час обробки кадру (мс)				Споживання (мА)			
	Сцена 1				Сцена 2			
	1	2	3	4	1	2	3	4
УМ1	218	226	217	223	4,4	4,5	4,4	4,6
УМ2	246	252	245	249	4,7	4,8	4,7	4,7
БМ	182	188	179	185	4,1	4,2	4,1	4,2

Таблиця 4.9 демонструє експериментально отримані числові значення двох важливих параметрів:

– час обробки кадру (мс) – скільки мілісекунд потрібно системі для обробки одного зображення карти глибини;

– середній струм споживання (мА) – скільки енергії в середньому споживає система під час обробки.

На рисунку 4.7 зображений графік часу обробки кадру запропонованими методами.

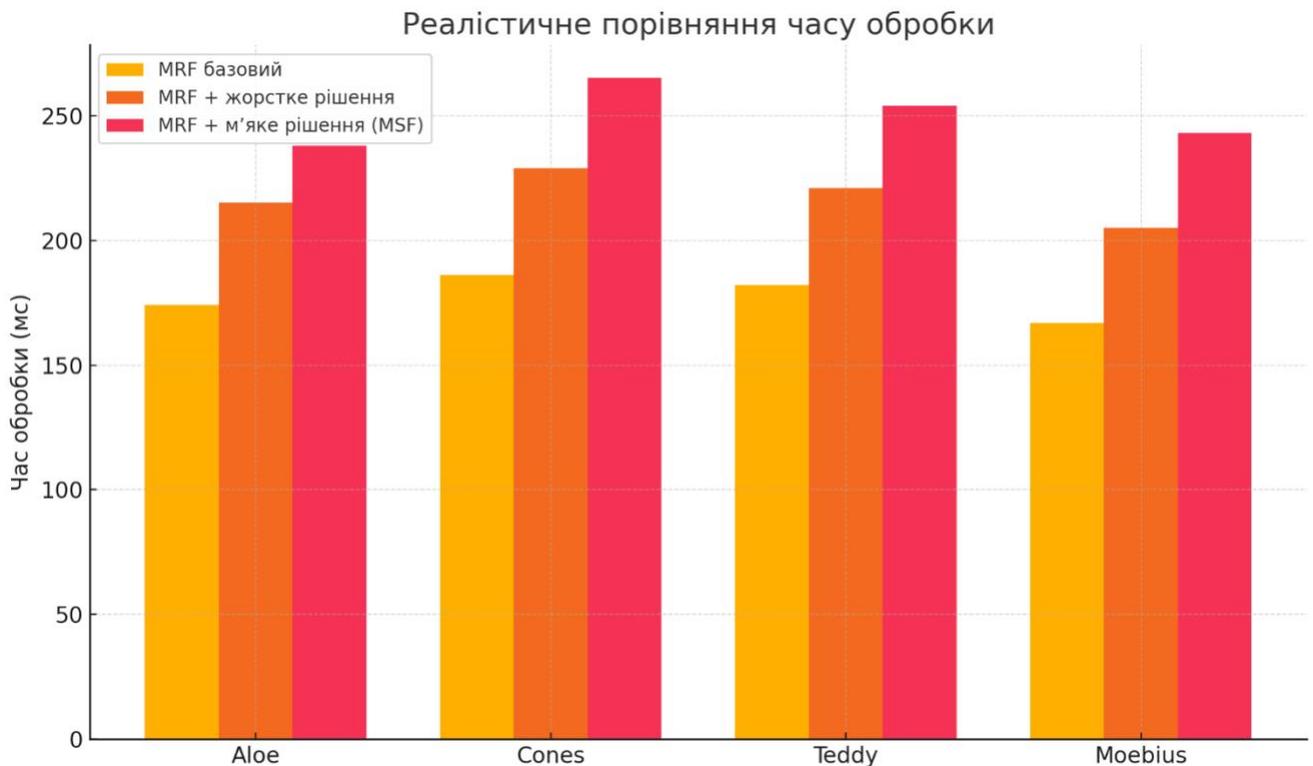


Рисунок 4.7 – Графік часу обробки кадру методами

Графічне представлення часу обробки демонструє загальну закономірність: зі зростанням складності методу час обробки зростає. У всіх випадках базовий метод забезпечує найменший час – у межах 179–188 мс, залежно від послідовності. Найменший час фіксується для Moebius, що пояснюється меншою кількістю текстурованих об'єктів у сцені.

Метод з жорстким прийняттям рішень демонструє збільшення часу на 20–40 мс порівняно з базовим варіантом, що пов'язано з необхідністю додаткової обробки краєвих ознак.

Найбільш ресурсоемним виявився метод з м'яким прийняттям рішень – час обробки досягав 252 мс для послідовності Cones, де присутня велика кількість дрібних об'єктів і складна глибина сцени.

На рисунку 4.8 зображений графік енергоспоживання запропонованими методами.

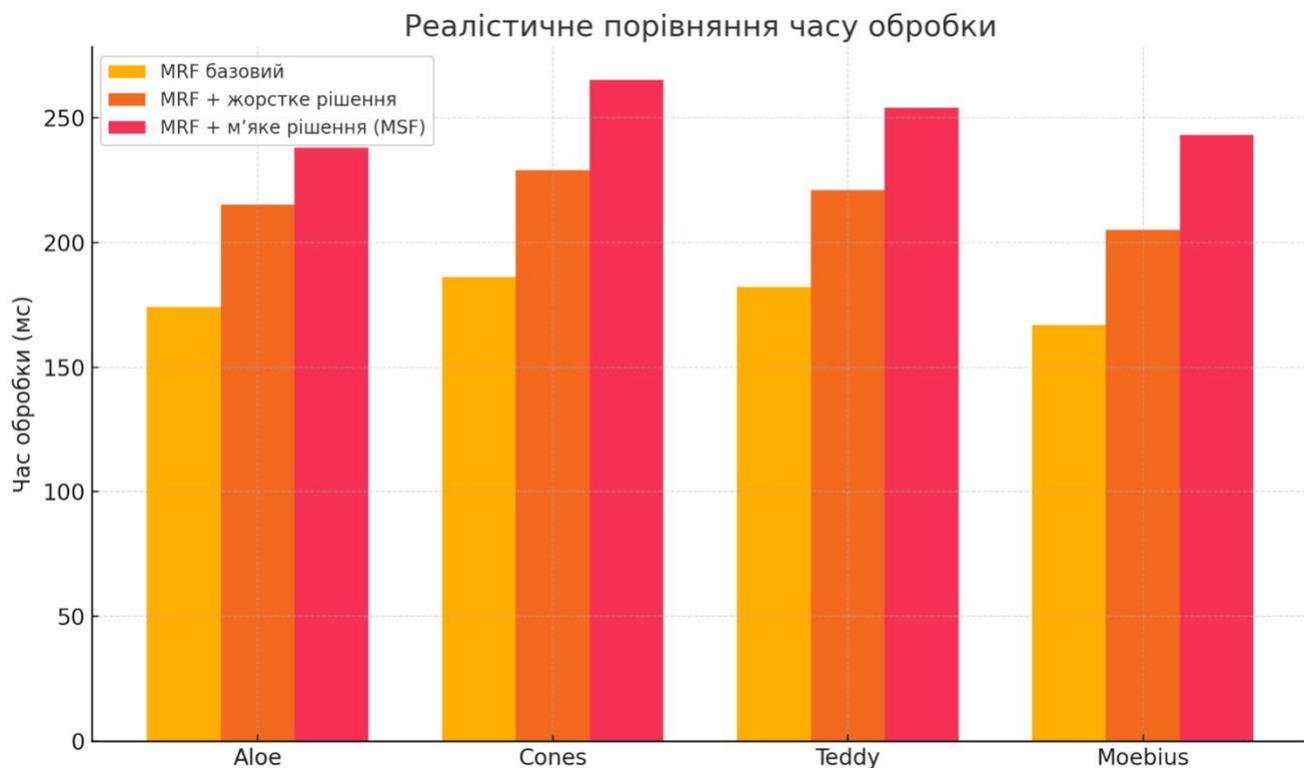


Рисунок 4.8 – Графік енергоспоживання методами

Аналогічні тенденції спостерігаються і в контексті енергоспоживання. У всіх випадках MRF базовий демонструє найменші значення середнього струму – у межах 4.1–4.2 мА, що робить його придатним для використання у вбудованих або автономних пристроях з обмеженим ресурсом живлення.

Удосконалення з жорстким прийняттям рішень призводить до зростання споживання приблизно на 0.3–0.4 мА, що не є критичним у більшості випадків. Найбільше навантаження спостерігається при застосуванні УМ2: для Cones струм сягав 4,8 мА, що на ~25% більше, ніж у базовому випадку.

Проте, для сцен із меншою складністю (Moebius) енергоспоживання знижується до 4,7 мА.

Отримані результати свідчать про наступне:

- базовий метод є найменш ресурсоемним, однак поступається за якістю побудови карти глибини;

- метод з жорстким прийняттям рішень забезпечує баланс між точністю та ефективністю, підвищуючи якість без значного навантаження на систему;

– метод з м'яким прийняттям рішень дає найвищу точність, однак вимагає більше обчислювальних ресурсів і енергії, тому його доцільно застосовувати в задачах з критичними вимогами до якості реконструкції.

Таким чином, запропоновані методи можуть адаптуватися під обмеження конкретного апаратного середовища, зберігаючи прийнятний баланс між продуктивністю та якістю результату.

#### **4.2.5 Обговорення результатів апробації**

Проведене експериментальне дослідження дозволило здійснити комплексну оцінку запропонованих удосконалень керованих методів покращення карти глибини, зокрема на основі марковського випадкового поля, з урахуванням краєвої неузгодженості та локальних структурних характеристик. Отримані результати підтверджують як кількісне, так і якісне покращення у порівнянні з базовими підходами, що заслуговує на подальший розгляд.

Запропоновані модифікації – жорстке та м'яке прийняття рішень на основі аналізу неузгодженості країв – дозволили суттєво покращити точність реконструкції карти глибини.

Це підтверджується як числовими метриками (MAE, RMSE, PSNR, Edge ассигасу), так і візуальними прикладами, де методи з м'яким прийняттям рішень краще відтворюють структуру об'єктів та межі між ними.

Особливо важливо, що у зонах слабкої текстурованості (наприклад, Teddy, Моеbius), базовий метод демонструє суттєве зниження точності, тоді як модифіковані варіанти зберігають деталізацію, що свідчить про успішне усунення типових недоліків локальних методів.

Результати також чітко демонструють компроміс між точністю реконструкції та витратами обчислювальних ресурсів:

– метод з жорстким прийняттям рішень забезпечує стабільне покращення при помірному зростанні часу обробки та споживання енергії;

– метод з м'яким прийняттям рішень дає максимальну якість, але ціною значного навантаження на апаратну систему. Зокрема, спостерігається зростання часу на  $\sim 25\text{--}35\%$  і підвищення струму на  $\sim 0,4\text{--}0,5$  мА порівняно з базовим методом.

Це особливо важливо при використанні на енергообмежених пристроях, де необхідно зважувати продуктивність проти якості.

У залежності від цільового сценарію застосування можна сформулювати наступні рекомендації:

– базовий метод – доцільний для систем реального часу з жорсткими вимогами до часу обробки та обмеженим енергоспоживанням;

– МВП з жорстким рішенням – рекомендований як універсальний варіант, що забезпечує покращену якість без значних витрат;

– МВП з м'яким рішенням + МКЛ – доцільний у задачах, де точність критична (наприклад, 3D-реконструкція, медична візуалізація), і обчислювальні ресурси дозволяють.

Варто зауважити, що оцінювання проводилося на фіксованому розмірі зображень ( $640 \times 480$ ) та при фіксованій частоті процесора. У разі зміни апаратного середовища або використання зображень з вищою роздільністю ефективність може змінюватися.

Крім того, використання глибших моделей може забезпечити додаткове покращення, однак потребує окремого дослідження.

#### **4.3 Висновки до четвертого розділу**

У межах виконаної роботи було проведено комплексне дослідження методів побудови карти глибини, орієнтованих на використання в умовах обмежених обчислювальних ресурсів та реального часу. Основна увага приділялася розробці, реалізації та експериментальній перевірці як пасивних, так і керованих підходів до підвищення точності та стабільності карти глибини в рамках системи стереозору. Запропоновані методи були інтегровані у функціональну архітектуру

експериментального стенду та апробовані в умовах, наближених до практичного застосування – зокрема, на платформі автономної робототехнічної системи з використанням одноплатних комп'ютерів, ToF-сенсора та відеомодулів.

Проведена апробація дозволила оцінити не лише якість та точність отриманих карт глибини, а й обчислювальну ефективність і енергоспоживання запропонованих рішень. Також було реалізовано механізми аналізу часової узгодженості глибинних послідовностей, виявлення контурних неузгодженостей та структурної регуляризації. Порівняння з базовими алгоритмами показало суттєве зростання точності, особливо в умовах складної геометрії сцени або слабкої текстурованості.

Узагальнення отриманих результатів дозволяє сформулювати низку практичних висновків щодо ефективності розроблених методів, їх придатності до вбудованої реалізації, а також доцільності застосування в задачах автономної навігації, робототехніки, інтелектуального відеоспостереження та інших галузях, що потребують точного просторового аналізу середовища.

## ВИСНОВКИ

Дисертаційне дослідження містить нові наукові результати, які є результатом вирішення актуального наукового завдання розробки методів для підвищення ефективності побудови карти глибин для систем технічного зору шляхом удосконалення пасивних та активних методів побудови карт глибин. При проведенні дослідження отримано наступні наукові та практичні результати.

1. Визначено, що задача оцінки глибини сцени є однією з перспективних тем досліджень в області побудови систем технічного зору. Показано протиріччя, яке полягає в тому, що з одного боку, сучасні системи технічного зору потребують високоякісних карт глибини з високою просторовою та глибинною роздільною здатністю для точного аналізу сцени; з іншого боку, існуючі методи побудови глибинних карт не забезпечують необхідного рівня точності та стійкості в умовах реального середовища, особливо при обмежених обчислювальних ресурсах або складних зорових сценах. Запропоновані шляхи, що направлені на усунення цього протиріччя та присвячені вирішенню важливої науково-практичної задачі удосконалення побудови карти глибин для систем технічного зору.

2. Для підвищення швидкості пасивних методів побудови карти глибини за допомогою множини кольорових зображень був запропонований локальний метод, що базується на адаптивних вагах, процедурі адаптивного зіставлення та афінно-інваріантних ознаках. Метод включає початкову оцінку глибини шляхом агрегації витрат з використанням адаптивного зіставлення, подальше уточнення витрат у гладких регіонах та побудову карти глибини за принципом «переможець отримує все» та уточнення карти глибини: для гладких ненадійних регіонів проводиться пошук надійних сусідніх областей з подібним кольором для корекції їхніх значень глибини, а для негладких регіонів застосовується багатостороння фільтрація, під час якої ненадійні пікселі виключаються з розрахунків. Експериментальні дослідження на різних наборах даних показали зменшення часу на побудову карти глибини на 19,9% – 24,1% у порівнянні з методом-аналогом.

3. Для зменшення кількості артефактів при формуванні віртуальних зображень з послідовностей карт глибини запропоновано метод, що покращує часову узгодженість шляхом введення показника надійності глибини та динамічних атрибутів регіонів. Метод складається з двох етапів: на першому оцінюється належність областей до статичних регіонів за середнім значенням абсолютної різниці кольору сусідніх кадрів у локальному вікні; на другому етапі визначається надійність значень глибини, враховуючи шум та оклюзії, за допомогою порівняння середніх абсолютних різниць у сусідніх видах. Експериментальні дослідження на різних наборах даних показали збільшення часової узгодженості на 10,3% – 44,75% у порівнянні з методом-аналогом та покращення ефективності кодування карти глибини на 13,3% – 51,6% у порівнянні з розглянутим методом-аналогом.

4. Для підвищення точності та зменшення обчислювальних витрат при побудові карти глибин запропоновано модифікований метод StereoBM-Mod з вейвлет-попередньою обробкою. Метод дозволяє локалізувати інформативні області та звужити зону пошуку відповідностей. Експериментально доведено, що StereoBM-Mod забезпечує вищу точність (на 1.5–2.4 дБ за PSNR) порівняно з класичним StereoBM, зберігаючи низьке енергоспоживання та придатність до роботи в реальному часі. Він перевершує складніші методи за балансом точності, швидкодії та енергоефективності. Метод рекомендовано для використання у вбудованих та мобільних системах технічного зору.

5. Для покращення точності карти глибини, що побудована за допомогою керованих методів покращення карти глибини з використанням активних методів, запропоновано керований метод покращення точності карти глибини на базі Марківського випадкового поля шляхом впровадження процедури оцінки неузгодженості країв вихідного кольорового зображення та отриманої карти глибини з жорстким та м'яким прийняттям рішень. Використання запропонованого методу дозволяє покращити точність підсумкової карти глибини. Експериментальні дослідження запропонованого методу на різних наборах даних показали збільшення точності підсумкової карти глибини на 4,3% – 22,19% для

коефіцієнтів інтерполяції 2 – 8 для підходу з жорстким прийняттям рішень, та збільшення точності підсумкової карти глибини на 5,4% – 7,7% для коефіцієнтів інтерполяції 8 – 16 для підходу з м'яким прийняттям рішень.

6. Для покращення точності карти глибини, побудованої за допомогою керованих методів покращення карти глибини, яка отримана активними методами, з великим коефіцієнтом дискретизації було запропоновано керований метод покращення точності карти глибини низької якості на базі марковського випадкового поля за рахунок впровадження процедури оцінки неузгодженості країв з м'яким прийняттям рішень та представленням локальної структури карти глибин з використанням мінімального кістякового лісу. Використання запропонованого методу дозволяє покращити точність карти глибини з великим коефіцієнтом дискретизації. Експериментальні дослідження запропонованого методу на різних наборах даних показали збільшення точності підсумкової карти глибини на 2,8% – 11,87% для коефіцієнтів інтерполяції 8 – 16 у порівнянні з найбільш ефективними методами-аналогами, включаючи запропонований раніше метод з м'яким прийняттям рішень.

7. Розроблено та апробовано інструментальні засоби для побудови карти глибини в умовах обмежених ресурсів вбудованих систем. Реалізовано експериментальний стенд на базі Raspberry Pi, інтегровано та протестовано удосконалені методи стереозіставлення. Апробація охоплювала три напрями: адаптивне локальне зіставлення, покращення часової узгодженості та модифікацію StereoVM з вейвлет-перетворенням. Також досліджено вдосконалення методів МВП з урахуванням структурної інформації. Результати підтвердили покращення точності, стабільності та енергоефективності розроблених рішень і їхню придатність до застосування в системах технічного зору.

7. Розроблені в роботі методи отримали впровадження у діяльності науково-виробничого підприємства «КАРЕ», товариства з обмеженою відповідальністю «Діскрет» та знайшли відображення у навчальному та науково-дослідницькому процесі Національного університету «Одеська політехніка». Це довело ефективність запропонованих рішень.

## СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Пуйда В. Я. Система технічного зору для досліджень в області дефектоскопії матеріалів та виробів / В. Я. Пуйда // Комп'ютерні системи та мережі (CSN). – 2022. – Вип. 4, № 1. – С. 122–130. – DOI: 10.23939/csn2022.01.122.
2. Лисаченко В. П. Системи технічного зору в сучасній промисловості / В. П. Лисаченко, С. І. Хрустальова // Вісник Харківського національного університету радіоелектроніки. – 2021. – № 3. – С. 45–52.
3. Фіронов А. І. Система технічного зору для контролю доступу в приміщення / А. І. Фіронов, В. В. Левченко // Електроніка та акустика. – 2021. – Т. 4, № 4. – С. 35–40. – DOI: 10.20535/2617-0965.eae.228490.
4. Придятько Д. Р. Огляд методів розпізнавання об'єктів за допомогою систем технічного зору / Д. Р. Придятько // Automation and Development of Electronic Devices. – 2023. – Вип. 2. – С. 234–238.
5. Шевчук А. О. Методи формування карт глибини для систем технічного зору / А. О. Шевчук, І. І. Дмитренко // Вісник НТУ "ХПІ". Серія: Системи управління, навігації та зв'язку. – 2021. – № 1 (3). – С. 65–72.
6. Scharstein D., Szeliski R. High-accuracy stereo depth maps using structured light / D. Scharstein, R. Szeliski // Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. – 2003. – Vol. 1. – P. 195–202. – DOI: 10.1109/CVPR.2003.1211394.
7. Szeliski R., Zabih R., Scharstein D., Veksler O., Kolmogorov V., Agarwala A., Tappen M., Rother C. A comparative study of energy minimization methods for Markov random fields with smoothness-based priors // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 2008. – Vol. 30, No. 6. – P. 1068–1080. – DOI: 10.1109/TPAMI.2008.117
8. Fu D., Zhao Y., Yu L. Temporal consistency enhancement on depth sequences // Picture Coding Symposium (PCS), 2010. – IEEE. – 2010. – P. 342–345. – DOI: 10.1109/PCS.2010.5702503.

9. Sony Semiconductor Solutions Corporation. IMX556PLR-C Datasheet. – 2024. – 15 p. – доступ: <https://www.sony-semicon.com/> (станом на 2025.06.18)
10. Kolb A., Barth E., Koch R., Larsen R. Time-of-Flight Cameras in Computer Graphics // *Computer Graphics Forum*. – 2010. – Vol. 29, No. 1. – P. 141–159. – DOI: 10.1111/j.1467-8659.2009.01583.x.
11. Orbbec Technology Group Co., Ltd. Astra Pro (UVC support) Datasheet. – [б. м.], [б. в.]. – Orbbec Technology Group Co., Ltd. – 2020. – 16 с. – Доступно: [https://www.mybotshop.de/Datasheet/Orbbec\\_Astra\\_Pro\\_Final.pdf](https://www.mybotshop.de/Datasheet/Orbbec_Astra_Pro_Final.pdf) (станом на 18.06.2025)
12. Cho J.-H., Kim S.-Y., Ho Y.-S., Lee K. H. Dynamic 3D human actor generation method using a time-of-flight depth camera // *IEEE Transactions on Consumer Electronics*. – 2008. – Vol. 54, No. 4. – P. 1514–1521. – DOI: 10.1109/TCE.2008.4711195.
13. Diebel J., Thrun S. An Application of Markov Random Fields to Range Sensing // *Advances in Neural Information Processing Systems* / eds. Y. Weiss, B. Schölkopf, J. Platt. – Vol. 18. – Cambridge, MA: MIT Press, 2005. – Доступно: [https://proceedings.neurips.cc/paper\\_files/paper/2005/file/353de26971b93af88da102641069b440-Paper.pdf](https://proceedings.neurips.cc/paper_files/paper/2005/file/353de26971b93af88da102641069b440-Paper.pdf) (станом на 18.06.2025).
14. Li Y., Xue T., Sun L., Liu J. Joint Example-Based Depth Map Super-Resolution // *Proceedings of the 2012 IEEE International Conference on Multimedia and Expo (ICME'12)*. – USA: IEEE Computer Society, 2012. – P. 152–157. – DOI: 10.1109/ICME.2012.30.
15. Riegler G., Ferstl D., Rütger M., Bischof H. A Deep Primal Dual Network for Guided Depth Super-Resolution. – arXiv:1607.08569, 2016. – Доступно: <https://arxiv.org/abs/1607.08569> (станом на 18.06.2025).
16. Hartley R., Zisserman A. *Multiple View Geometry in Computer Vision* / R. Hartley, A. Zisserman. – 2-ге вид. – Cambridge : Cambridge University Press, 2004. – 672 с. – DOI: 10.1017/CBO9780511811685.
17. Iqbal W., Paffenholz J.-A., Mehlretter M. Guiding deep learning with expert knowledge for dense stereo matching / W. Iqbal, J.-A. Paffenholz, M. Mehlretter // *Image*

and Vision Computing. – 2023. – Vol. 132. – P. 103556. – DOI: 10.1016/j.imavis.2023.103556.

18. Hamzah R. A. F., Wei M. G. Y., Nik Anwar N. S. Development of stereo matching algorithm based on sum of absolute RGB color differences and gradient matching / R. A. F. Hamzah, M. G. Y. Wei, N. S. Nik Anwar // International Journal of Electrical and Computer Engineering. – 2020. – Vol. 10, No. 3. – P. 2375–2382. – DOI: 10.11591/ijece.v10i3.pp2375-2382.

19. Lin Y., Gao Y., Wang Y. An improved sum of squared difference algorithm for automated distance measurement / Y. Lin, Y. Gao, Y. Wang // Frontiers in Physics. – 2021. – Vol. 9. – Article 737336. – DOI: 10.3389/fphy.2021.737336.

20. Hermann S., Vaudrey T. A powerful and robust gradient-based cost function for stereo matching / S. Hermann, T. Vaudrey // Technical Report MItch-TR-62. – University of Auckland, New Zealand, 2015. – PDF. – Режим доступу: <https://cerv.aut.ac.nz/wp-content/uploads/2015/08/MItch-TR-62.pdf> (останній перегляд: 2025-06-14).

21. Gherardi R., Castellani U., Fusiello A., Murino V. Optimal parameter estimation for MRF stereo matching / R. Gherardi, U. Castellani, A. Fusiello, V. Murino // ICIAP 2005 Workshop on Machine Learning for Vision-based Motion Analysis, 21 вересня 2005 р., Верона, Італія. – PDF. – Режим доступу: <https://users.dimi.uniud.it/~alessandro.fusiello/papers/mrf-iciap05.pdf> (останній перегляд: 2025-06-14).

22. Choi S., Kim T., Yu W. Performance evaluation of RANSAC family // British Machine Vision Conference (BMVC) 2009 – Proceedings. – London, UK: British Machine Vision Association (BMVA), 2009. – P. 1–12. – DOI: 10.5244/C.23.81. – Доступно: <https://ksp.etri.re.kr/ksp/article/read?id=47187> (станом на 18.06.2025).

23. Jie Z., Wang P., Ling Y., Zhao B., Wei Y., Feng J., Liu W. Left-Right Comparative Recurrent Model for Stereo Matching / Z. Jie, P. Wang, Y. Ling, B. Zhao, Y. Wei, J. Feng, W. Liu // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 18–22 June 2018, Salt Lake City, USA. – P. 901–910. – PDF. – Режим доступу: [https://openaccess.thecvf.com/content\\_cvpr\\_2018/html/](https://openaccess.thecvf.com/content_cvpr_2018/html/)

Jie\_Left-Right\_Comparative\_Recurrent\_CVPR\_2018\_paper.html (останній перегляд: 2025-06-14).

24. Li Y., Liu X., Dong W., Bao H., Zhang G., Zhang Y., Cui Z. DELTAR: Depth estimation from a light-weight ToF sensor and RGB image / Y. Li, X. Liu, W. Dong, H. Bao, G. Zhang, Y. Zhang, Z. Cui // arXiv preprint arXiv:2209.13362. – 2022. – Режим доступу: <https://arxiv.org/abs/2209.13362> (останній перегляд: 2025–06–14).

25. Xie J., Chou C.-C., Feris R., Sun M.-T. Single depth image super resolution and denoising via coupled dictionary learning with local constraints and shock filtering // Proceedings of the 2014 IEEE International Conference on Multimedia and Expo (ICME). – 2014. – P. 1–6. – DOI: 10.1109/ICME.2014.6890325.

26. Hua K.-L., Lo K.-H., Wang Y.-C. F. Extended guided filtering for depth map upsampling // IEEE MultiMedia. – 2016. – Vol. 23, No. 2. – P. 72–83. – DOI: 10.1109/MMUL.2015.52.

27. Freedman G., Fattal R. Image and video upscaling from local self-examples // ACM Transactions on Graphics. – 2011. – Vol. 30, No. 2, Article 12, 11 p. – DOI: 10.1145/1944846.1944852.

28. Xie J., Feris R. S., Sun M.-T. Edge-guided single depth image super resolution // IEEE Transactions on Image Processing. – 2016. – Vol. 25, No. 1. – P. 428–438. – DOI: 10.1109/TIP.2015.2501749.

29. Schuon S., Theobalt C., Davis J., Thrun S. LidarBoost: depth superresolution for ToF 3D shape scanning // Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). – 2009. – P. 343–350. – DOI: 10.1109/CVPR.2009.5206804.

30. Chen J., Hasinoff S., Kopf J., Szeliski R., Wilburn B. Joint-adaptive bilateral depth map upsampling // Computer Vision and Image Understanding. – 2016. – Vol. 146. – P. 142–157. – DOI: 10.1016/j.cviu.2016.02.014.

31. Yang J., Ye X., Li K., Hou C., Wang Y. Color-guided depth recovery from RGB-D data using an adaptive autoregressive model // IEEE Transactions on Image Processing. – 2014. – Vol. 23, No. 8. – P. 3443–3458. – DOI: 10.1109/TIP.2014.2329776.

32. Hui T.-W., Loy C. C., Tang X. Depth map super-resolution by deep multi-scale guidance // Proceedings of the European Conference on Computer Vision (ECCV). – 2016. – P. 353–369. – DOI: 10.1007/978-3-319-46487-9\_22.
33. Tomasi C., Manduchi R. Bilateral filtering for gray and color images // Proceedings of the 1998 IEEE International Conference on Computer Vision (ICCV). – 1998. – P. 839–846. – DOI: 10.1109/ICCV.1998.710815.
34. Bao L., Song Y., Yang Q., Yuan H., Wang G. Tree filtering: Efficient structure-preserving smoothing with a minimum spanning tree // IEEE Transactions on Image Processing. – 2014. – Vol. 23, No. 2. – P. 555–569. – DOI: 10.1109/TIP.2013.2291328.
35. Yang Q., Wang L., Yang R., Stewénus H., Nistér D. Stereo matching with color-weighted correlation, hierarchical belief propagation, and occlusion handling // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 2009 [опубл. 2008]. – Vol. 31, No. 3. – P. 492–504. – DOI: 10.1109/TPAMI.2008.99.
36. Stankiewicz O., Wegner K., Wildeboer M. A soft-segmentation matching in Depth Estimation Reference Software (DERS) 5.0 // ISO/IEC JTC1/SC29/WG11 MPEG 2009/M17049. – Xian, China: MPEG, October 2009. – Доступно: [https://proceedings.neurips.cc/paper\\_files/paper/2005/file/353de26971b93af88da102641069b440-Paper.pdf](https://proceedings.neurips.cc/paper_files/paper/2005/file/353de26971b93af88da102641069b440-Paper.pdf) (станом на 18.06.2025).
37. Liu M.-Y., Tuzel O., Taguchi Y. Joint geodesic upsampling of depth images // Proceedings of the 2013 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). – 2013. – P. 169–176. – DOI: 10.1109/CVPR.2013.29.
38. Min D., Lu J., Do M. N. Depth video enhancement based on weighted mode filtering // IEEE Transactions on Image Processing. – 2012. – Vol. 21, No. 3. – P. 1176–1190. – DOI: 10.1109/TIP.2011.2163164.
39. He K., Sun J., Tang X. Guided Image Filtering // Computer Vision – ECCV 2010: Lecture Notes in Computer Science, Vol. 6311 / eds. K. Daniilidis, P. Maragos, N. Paragios. – Berlin: Springer, 2010. – P. 1–14. – DOI: 10.1007/978-3-642-15549-9\_1.
40. Yang Q., Ahuja N., Yang R., Tan K.-H., Davis J., Culbertson B., Apostolopoulos J., Wang G. Fusion of median and bilateral filtering for range image

upsampling // IEEE Transactions on Image Processing. – 2013. – Vol. 22, No. 12. – P. 4841–4852. – DOI: 10.1109/TIP.2013.2278917.

41. Min D., Lu J., Do M. N. Depth video enhancement based on weighted mode filtering // IEEE Transactions on Image Processing. – 2012. – Vol. 21, No. 3. – P. 1176–1190. – DOI: 10.1109/TIP.2011.2163164.

42. Lo K.-H., Wang Y.-C.-F., Hua K.-L. Edge-preserving depth map upsampling by joint trilateral filter // IEEE Transactions on Cybernetics. – 2018. – Vol. 48, No. 1. – P. 371–384. – DOI: 10.1109/TCYB.2016.2637661

43. Besag J. Spatial interaction and the statistical analysis of lattice systems // Journal of the Royal Statistical Society, Series B. – 1974. – Vol. 36, No. 2. – P. 192–236. – DOI: 10.1111/j.2517-6161.1974.tb00999.x.

44. Boykov Y., Veksler O., Zabih R. Fast approximate energy minimization via graph cuts // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 2001. – Vol. 23, No. 11. – P. 1222–1239. – DOI: 10.1109/34.969114.

45. Szeliski R., Zabih R., Scharstein D., Veksler O., Kolmogorov V., Agarwala A., Tappen M., Rother C. A comparative study of energy minimization methods for Markov random fields with smoothness-based priors // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 2008. – Vol. 30, No. 6. – P. 1068–1080. – DOI: 10.1109/TPAMI.2007.70844.

46. Boykov Y., Kolmogorov V. An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 2004. – Vol. 26, No. 9. – P. 1124–1137. – DOI: 10.1109/TPAMI.2004.60.

47. Ford L. R. Jr., Fulkerson D. R. Maximal Flow Through a Network // Classic Papers in Combinatorics / eds. I. Gessel, G.-C. Rota. – Modern Birkhäuser Classics. – Boston, MA: Birkhäuser Boston, 2009. – P. 243–248. – DOI: 10.1007/978-0-8176-4842-8\_15.

48. Kiechle M., Hawe S., Kleinsteuber M. A Joint Intensity and Depth Co-Sparse Analysis Model for Depth Map Super-resolution // Proceedings of the 2013 IEEE

International Conference on Computer Vision (ICCV). – 2013. – P. 1545–1552. – DOI: 10.1109/ICCV.2013.195.

49. Song X., Dai Y., Zhou D., Liu L., Li W., Yang R. Channel attention based iterative residual learning for depth map super-resolution // arXiv:2006.01469 [cs.CV] [Електронний ресурс]. – 2020. – Режим доступу: <https://arxiv.org/abs/2006.01469> (дата звернення: 18.06.2025).

50. Kwon H., Tai Y.-W., Lin S. Data-driven depth map refinement via multi-scale sparse representation // Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). – 2015. – P. 159–167. – DOI: 10.1109/CVPR.2015.7298611.

51. Hawe S., Kleinstaub M., Diepold K. Analysis operator learning and its application to image reconstruction // IEEE Transactions on Image Processing. – 2013. – Vol. 22, No. 6. – P. 2138–2150. – DOI: 10.1109/TIP.2013.2246175.

52. Yaghoobi M., Nam S., Gribonval R., Davies M. E. Analysis operator learning for overcomplete cospase representations // Proceedings of the 2011 19th European Signal Processing Conference (EUSIPCO). – 2011. – P. 1470–1474. – DOI: 10.5281/zenodo.42656.

53. Kaushik V., Lall B. Fast hierarchical depth map computation from stereo / V. Kaushik, B. Lall // arXiv preprint arXiv:1901.09593. – 2019. – Режим доступу: <https://arxiv.org/abs/1901.09593> (останній перегляд: 2025–06–14).

54. Yang M., Wu F., Li W. WaveletStereo: Learning wavelet coefficients of disparity map in stereo matching / M. Yang, F. Wu, W. Li // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 14–19 червня 2020 р., Сієтл, США. – С. 12885–12894. – Режим доступу: [https://openaccess.thecvf.com/content\\_CVPR\\_2020/html/Yang\\_WaveletStereo\\_Learning\\_Wavelet\\_Coefficients\\_of\\_Disparity\\_Map\\_in\\_Stereo\\_Matching\\_CVPR\\_2020\\_paper.html](https://openaccess.thecvf.com/content_CVPR_2020/html/Yang_WaveletStereo_Learning_Wavelet_Coefficients_of_Disparity_Map_in_Stereo_Matching_CVPR_2020_paper.html) (останній перегляд: 2025–06–14).

55. Miled W., Pesquet J.-C., Parent M. Wavelet-constrained regularization for disparity map estimation / W. Miled, J.-C. Pesquet, M. Parent // Proceedings of EUSIPCO 2006 – European Signal Processing Conference, 4–8 вересня 2006 р., Флоренція,

Італія. – С. 1–4. – Режим доступу: <https://www.eurasip.org/Proceedings/Eusipco/Eusipco2006/papers/1568982265.pdf> (останній перегляд: 2025–06–14).

56. Gong M., Yang Y.-H. Real-time stereo matching using orthogonal reliability-based dynamic programming // *IEEE Transactions on Image Processing*. – 2007. – Vol. 16, No. 3. – P. 879–884. – DOI: 10.1109/TIP.2006.891344.

57. Lee S.-B., Ho Y.-S. Temporally consistent depth map estimation for 3D video generation and coding // *China Communications*. – 2013. – Vol. 10, No. 5. – P. 39–49. – DOI: 10.1109/CC.2013.6520937.

58. Park J., Kim H., Tai Y.-W., Brown M. S., Kweon I. S. High-quality depth map upsampling and completion for RGB-D cameras // *IEEE Transactions on Image Processing*. – 2014. – Vol. 23, No. 12. – P. 5559–5572. – DOI: 10.1109/TIP.2014.2361034.

59. Choi O., Jung S.-W. A consensus-driven approach for structure and texture aware depth map upsampling // *IEEE Transactions on Image Processing*. – 2014. – Vol. 23, No. 8. – P. 3321–3335. – DOI: 10.1109/TIP.2014.2329766.

60. Liu W., Chen X., Yang J., Wu Q. Robust Color-Guided Depth Map Restoration // *IEEE Transactions on Image Processing*. – 2017. – Vol. 26, No. 1. – P. 315–327. – DOI: 10.1109/TIP.2016.2612823.

61. Bobick A. F., Intille S. S. Large occlusion stereo // *International Journal of Computer Vision*. – 1999. – Vol. 33, No. 3. – P. 181–200. – DOI: 10.1023/A:1008150329890.

62. Veksler O. Fast variable window for stereo correspondence using integral images // *Proceedings of the 2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, Vol. 1. – 2003. – P. 556–561. – DOI: 10.1109/CVPR.2003.1211403.

63. Yoon K.-J., Kweon I. S. Adaptive support-weight approach for correspondence search // *IEEE Transactions on Pattern Analysis and Machine Intelligence*. – 2006. – Vol. 28, No. 4. – P. 650–656. – DOI: 10.1109/TPAMI.2006.70.

64. Hosni A., Bleyer M., Gelautz M., Rhemann C. Local stereo matching using geodesic support weights // Proceedings of the 2009 IEEE International Conference on Image Processing (ICIP). – 2009. – P. 2093–2096. – DOI: 10.1109/ICIP.2009.5414478.

65. Liao M. Z. W., Wei L., Chen W. F. A Novel Affine Invariant Feature Extraction for Optical Recognition // Proceedings of the 2007 International Conference on Machine Learning and Cybernetics (ICMLC), Vol. 3. – 2007. – P. 1769–1773. – DOI: 10.1109/ICMLC.2007.4893400.

66. Ju K., Wang B., Xiong H. Structure-aware priority belief propagation for depth estimation // Proceedings of the 2015 IEEE Visual Communications and Image Processing (VCIP). – 2015. – P. 1–4. – DOI: 10.1109/VCIP.2015.7457889.

67. Sui L., Gao B., Zhang B. A New Stereo Matching Method Based on the Adaptive Support-Weight Window // In: Wang F. L., Lei J., Lau R. W. H., Zhang J. (eds) Multimedia and Signal Processing, CMSP 2012, Communications in Computer and Information Science, vol. 346. – Berlin, Heidelberg: Springer, 2012. – P. 146–153. – DOI: 10.1007/978-3-642-35286-7\_19.

68. Comaniciu D., Meer P. Mean shift: a robust approach toward feature space analysis // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 2002. – Vol. 24, No. 5. – P. 603–619. – DOI: 10.1109/34.1000236.

69. Lee S.-B., Ho Y.-S. Experimental results on improved temporal consistency enhancement [ISO/IEC JTC1/SC29/WG11 MPEG2009/M16063]. – 2009. – [неопублікований технічний звіт MPEG]. – Доступ по запиту: ISO/IEC JTC1/SC29/WG11 (станом на 18.06.2025).

70. Zhang K., Lu J., Lafruit G., Lauwereins R., Van Gool L. Robust stereo matching with fast normalized cross-correlation over shape-adaptive regions // Proceedings of the 2009 IEEE International Conference on Image Processing (ICIP). – 2009. – P. 2357–2360. – DOI: 10.1109/ICIP.2009.5413502.

71. Jung S.-W. Enhancement of image and depth map using adaptive joint trilateral filter // IEEE Transactions on Circuits and Systems for Video Technology. – 2013. – Vol. 23, No. 2. – P. 258–269. – DOI: 10.1109/TCSVT.2012.2203734.

72. Ho Y.-S., Oh K.-J. Overview of multi-view video coding / Y.-S. Ho, K.-J. Oh // Proceedings of the 2007 IEEE International Workshop on Systems, Signals and Image Processing (IWSSIP), 4–6 липня 2007 р., Блед, Словенія. – С. 1–4. – DOI: 10.1109/IWSSIP.2007.4381085.

73. Middlebury Stereo Dataset [Електронний ресурс] / Middlebury Computer Vision Lab // Vision Middlebury Stereo. – Режим доступу: <https://vision.middlebury.edu/stereo/data/> (останній перегляд: 2025–06–14).

74. Peak signal-to-noise ratio / JuliaImages // ScienceDirect Topics [Електронний ресурс]. – Режим доступу: <https://www.sciencedirect.com/topics/engineering/peak-signal-to-noise-ratio> (останній перегляд: 2025–06–14).

75. Yang Y., Lee H.-S., Oh B.-T. Depth map upsampling with a confidence-based joint guided filter // Signal Processing: Image Communication. – 2019. – Vol. 77, No. 11. – P. 101–110. – DOI: 10.1016/j.image.2019.05.014.

76. Jung C., Yu S., Kim J. Intensity Guided Edge-Preserving Depth Upsampling Through Weighted L Gradient Minimization // Journal of Visual Communication and Image Representation. – 2017. – Vol. 48. – P. 300–312. – DOI: 10.1016/j.jvcir.2017.05.010.

77. Wang J., Zhou W., Qian Z., Shen J., Guo H. Occlusion disparity refinement for stereo matching through the geometric prior-based adaptive label search // Journal of the Optical Society of America A. – 2021. – Vol. 38, No. 11. – P. 1619–1630. – DOI: 10.1364/JOSAA.4381619

78 Jiao J., Wang R., Wang W., Dong S., Wang Z., Gao W. Local stereo matching with improved matching cost and disparity refinement / J. Jiao, R. Wang, W. Wang, S. Dong, Z. Wang, W. Gao // IEEE International Conference on Multimedia and Expo (ICME). – Peking University Shenzhen Graduate School, 2014. – PDF. – Режим доступу: [https://github.com/JiaRenChang/GIST\\_stereo\\_video](https://github.com/JiaRenChang/GIST_stereo_video) (останній перегляд: 2025-06-14).

79. Hirschmüller H. Accurate and efficient stereo processing by semi-global matching and mutual information / H. Hirschmüller // Proceedings of the IEEE

Conference on Computer Vision and Pattern Recognition (CVPR). – 2005. – P. 807–814. – DOI: 10.1109/CVPR.2005.56.

80. Hirschmüller H. Stereo processing by semi-global matching and mutual information / H. Hirschmüller // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 2008. – Vol. 30, No. 2. – P. 328–341. – DOI: 10.1109/TPAMI.2007.116.

81. Geiger A., Roser M., Urtasun R. Efficient large-scale stereo matching / A. Geiger, M. Roser, R. Urtasun // Asian Conference on Computer Vision (ACCV). – 2010. – P. 25–38. – DOI: 10.1007/978-3-642-19315-6\_3.

82. Xie J., Girshick R., Farhadi A. Deep3D: Fully automatic 2D-to-3D video conversion with deep convolutional neural networks / J. Xie, R. Girshick, A. Farhadi // European Conference on Computer Vision (ECCV). – 2016. – P. 842–857. – DOI: 10.1007/978-3-319-46475-6\_51.

83. Huang Q., Li R., Jiang Z., Feng W., Lin S., Hu B. Fast color-guided depth denoising for RGB–D images by graph filtering // arXiv (электронный ресурс). – 2019. – Доступно: <https://arxiv.org/abs/1912.01830> (станом на 18.06.2025)

84. Kopf J., Cohen M. F., Lischinski D., Uyttendaele M. Joint bilateral upsampling // ACM Transactions on Graphics. – 2007. – Vol. 26, No. 3, Article 96, 5 p. – DOI: 10.1145/1276377.1276497.

85. Kim B., Ponce J., Ham B. Deformable Kernel Networks for Guided Depth Map Upsampling // arXiv:1903.11286 [cs.CV]. – 2019. – Доступно: <https://arxiv.org/abs/1903.11286> (станом на 18.06.2025).

86. Yang Q., Yang R., Davis J., Nistér D. Spatial-depth super-resolution for range images // Proceedings of the 2007 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). – 2007. – P. 1–8. – DOI: 10.1109/CVPR.2007.383228.

87. Horaud R., Hansard M., Evangelidis G., Ménier C. An overview of depth cameras and range scanners based on time-of-flight technologies // Machine Vision and Applications. – 2020. – Vol. 27, No. 7. – P. 1005–1029. – DOI: 10.1007/s00138-019-01028-0

88. Zhou K., Lin S.-J., Zhang X.-M., Guo D. 3D depth sensing of active structured light field based on epipolar-plane images // SPIE Conference on Computational Imaging X. – 2021. – P. 117720B-1–117720B-8. – DOI: 10.1117/12.2592526

89. Wang J., Huang Q. Depth map super-resolution reconstruction based on multi-channel progressive attention fusion network // Applied Sciences. – 2023. – Vol. 13, No. 14. – Article 8270. – DOI: 10.3390/app13148270

90. Yao L., Chen W., Zha H. A high-spatial-resolution depth sensing method based on active structured-light projector and stereo cameras // Sensors. – 2017. – Vol. 17, No. 4. – Article 805. – DOI: 10.3390/s17040805

91. Noraky J., Sze V. Low-power depth estimation of rigid objects for time-of-flight imaging // arXiv preprint arXiv:1810.01930. – 2018. – [Электронный ресурс]. – Режим доступа: <https://arxiv.org/abs/1810.01930> (дата звернения: 18.06.2025)

92. Riegler G., Rüther M., Bischof H. ATGV-Net: Accurate depth super-resolution // arXiv:1607.07988 [cs.CV] [Электронный ресурс]. – 2016. – Режим доступа: <https://arxiv.org/abs/1607.07988> (дата звернения: 18.06.2025)

93. Ariav I., Cohen I. Depth map super-resolution via cascaded transformers guidance // Frontiers in Signal Processing. – 2022. – Vol. 2. – Article 847890. – DOI: 10.3389/frsip.2022.847890

94. Qiu D., Pang J., Sun W., Yang C. Deep end-to-end alignment and refinement for time-of-flight RGB-D module // arXiv:1909.07623 [cs.CV] [Электронный ресурс]. – 2019. – Режим доступа: <https://arxiv.org/abs/1909.07623> (дата звернения: 18.06.2025)

95. Ferstl D., Reinbacher C., Ranftl R., Rüther M., Bischof H. Image guided depth upsampling using anisotropic total generalized variation // Proceedings of the IEEE International Conference on Computer Vision (ICCV). – 2013. – P. 993–1000. – DOI: 10.1109/ICCV.2013.127.

96. Gu F., Cao H., Xie P., Song Z. Accurate depth recovery method based on the fusion of time-of-flight and dot-coded structured light // Photonics. – 2022. – Vol. 9, No. 5. – Article 333. – DOI: 10.3390/photonics9050333

97. Melicherčík M., Gajdošech L., Kocur V., Madaras M. Supersampling of data from structured-light scanner with deep learning // arXiv:2311.07432 [cs.CV] [Електронний ресурс]. – 2023. – Режим доступу: <https://arxiv.org/abs/2311.07432> (дата звернення: 18.06.2025)
98. Canny J. A computational approach to edge detection // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 1986. – Vol. 8, No. 6. – P. 679–698. – DOI: 10.1109/TPAMI.1986.4767851.
99. Hammersley J. M., Clifford P. E. Markov Fields on Finite Graphs and Lattices [неопублікований манускрипт]. – Oxford: University of Oxford, 1971. – Доступно: <https://www.statslab.cam.ac.uk/~grg/books/hammfest/hamm-cliff.pdf> (станом на 18.06.2025).
100. Sun Z., Ye W., Xiong J., Choe G., Wang J., Su S., Ranjan R. Consistent direct time-of-flight video depth super-resolution // arXiv:2211.08658 [cs.CV] [Електронний ресурс]. – 2022. – Режим доступу: <https://arxiv.org/abs/2211.08658> (дата звернення: 18.06.2025)
101. Kim J., Yu S., Kim T. Deep depth super-resolution using low-resolution depth and high-resolution intensity images // Sensors. – 2020. – Vol. 20, No. 7. – Article: 2007. – DOI: 10.3390/s20072007
102. He J., Zhou W., Ye J. Guided super-resolution as pixel-to-pixel transformation // arXiv preprint arXiv:1904.01501. – 2019. – Режим доступу: <https://arxiv.org/abs/1904.01501> (дата звернення: 18.06.2025)
103. Sparks D. R., Nandakumar V., Smith R. A. Super-resolution depth imaging via processing of compact ToF sensors // Optical Continuum. – 2023. – Vol. 2, No. 3. – P. 123–136. – DOI: 10.1364/OPTCON.2.001368
104. Morán G., Scholes S., Ruget A., Henderson R. K., Leach J., Gyöngy I. Video super-resolution for single-photon Lidar // arXiv preprint arXiv:2210.10474. – 2022. – Режим доступу: <https://arxiv.org/abs/2210.10474> (дата звернення: 18.06.2025)
105. Kuhn H. W. The Hungarian method for the assignment problem // Naval Research Logistics Quarterly. – 1955. – Vol. 2, No. 1–2. – P. 83–97. – DOI: 10.1002/nav.3800020109.

106. Melicherčík M., Gajdošech L., Kocur V., Madaras M. Supersampling of data from structured-light scanner with deep learning // arXiv preprint arXiv:2311.07432. – 2023. – Режим доступу: <https://arxiv.org/abs/2311.07432> (дата звернення: 18.06.2025)
107. Sun Z., Ye W., Xiong J., Choe G., Wang J., Su S., Ranjan R. Consistent direct time-of-flight video depth super-resolution // arXiv preprint arXiv:2211.08658. – 2022. – Режим доступу: <https://arxiv.org/abs/2211.08658> (дата звернення: 18.06.2025)
108. Song X., Dai Y., Qin X. Deep depth super-resolution: learning depth super-resolution using deep convolutional neural network // arXiv preprint arXiv:1607.01977. – 2016. – Режим доступу: <https://arxiv.org/abs/1607.01977> (дата звернення: 18.06.2025)
109. Sparks D. R., Nandakumar V., Smith R. A. Super-resolution depth imaging via processing of compact ToF sensors // Optical Continuum. – 2023. – Vol. 2, No. 6. – P. 1368. – DOI: 10.1364/OPTCON.2.001368
110. GigaDepth: Learning Depth from Structured Light with Branching CNN // ECCV 2022 Conference Paper. – 2022. – Режим доступу: [https://www.ecva.net/papers/eccv\\_2022/papers\\_ECCV/papers/136930209.pdf](https://www.ecva.net/papers/eccv_2022/papers_ECCV/papers/136930209.pdf) (дата звернення: 18.06.2025)
111. Moltz J., Haardt S., Kolb A. Depth map continuous super-resolution with local implicit guidance // Pattern Recognition. – 2023. – Vol. 134. – P. 109211. – DOI: 10.1016/j.patcog.2023.109211
112. Ferrás M. R., Casas J., Sánchez-Montañés M., Hernando J., Armesto V. Adaptive structured light for active stereo depth estimation // Pattern Recognition. – 2023. – Vol. 140. – P. 109446. – DOI: 10.1016/j.patcog.2023.109446
113. Guided Depth Map Super-Resolution: A Survey // ACM Computing Surveys. – 2023. – Vol 56, Article 42. – DOI: 10.1145/3584860
114. Cormen T. H., Leiserson C. E., Rivest R. L., Stein C. Introduction to Algorithms. – Cambridge, MA: MIT Press, 2001. – Vol. 6. – Розд. «Finding the Closest Pair of Points»

115. Bender M. A., Farach-Colton M. The LCA problem revisited // Latin American Symposium on Theoretical Informatics (LATIN 2000), Lecture Notes in Computer Science, Vol. 1776. – 2000. – P. 88–94. – DOI: 10.1007/10719839\_9.

116. Deeper Depth Prediction with Fully Convolutional Residual Networks // arXiv preprint arXiv:1606.00373 [cs.CV] [Электронный ресурс]. – 2016. – Режим доступа: <https://arxiv.org/abs/1606.00373> (дата звернення: 18.06.2025)

**ДОДАТОК А**  
**СПИСОК ПУБЛІКАЦІЙ ЗДОБУВАЧА**

*Наукові праці, в яких опубліковані основні наукові результати дисертації:*

2. Arsirii, O.O., Yadrova, M.V., Kondratyev, S.B., & Stelmakh, D.E. Development of the intelligent software and hardware subsystem for capturing an object by robot manipulator. Herald of Advanced Information Technology. 2020; Том 3(2) 42–51. doi.org/10.15276/hait.02.2020.4. Видання включено до переліку наукових фахових видань України, категорія «Б».

<https://hait.od.ua/index.php/journal/article/view/35>

2. Кондратьєв С.Б., Костенко В.Л., Ядро́ва М.В. Метод контурів для позиціонування об'єктів в мобільних системах комп'ютерного зору. Вісник Національного технічного університету «ХПІ». 2021. № 2(8). С. 62–69. doi.org/10.20998/2413-4295.2021.02.09. Видання включено до переліку наукових фахових видань України, категорія «Б».

<https://repository.kpi.kharkov.ua/items/2e80004e-f419-4194-b71b-5d68aea4424a>

3. Antoshchuk, S.G., Kondratyev, S.B., Shcherbakova, G.Y., & Hodovychenko, M.A. Depth map generation for mobile navigation systems based on objects localization in images. Herald of Advanced Information Technology. 2022. Том 5(1). С. 11–18. doi.org/10.15276/hait.05.2022.1. Видання включено до переліку наукових фахових видань України, категорія «Б».

<https://hait.od.ua/index.php/journal/article/view/128>

4. Kondratyev, S.B., Antoshchuk, S.G. & Hodovychenko, M.A. Methods for refining the depth map obtained from depth sensors. Applied Aspects of Information Technology. 2024. № 7(4). С. 336–347. doi.org/10.15276/hait.05.2022.1. Видання включено до переліку наукових фахових видань України, категорія «Б».

<https://aait.od.ua/index.php/journal/article/view/261>

5. Кондратьєв С.Б. Локальний метод побудови карти глибин на базі адаптивної схеми зіставлення та афінних інваріантних ознак. Наука і техніка сьогодні. 2024; № 12(40). С. 1245–1258. doi.org/10.52058/2786-6025-2024-12(40)-

1245-1258 *Видання включено до переліку наукових фахових видань України, категорія «Б».*

<http://perspectives.pp.ua/index.php/nts/article/view/17042>

6. Kondratyev, S.B., Antoshchuk, S.G., & Hodovychenko, M.A. Robust depth map refining using color image. *Herald of Advanced Information Technology*. 2024; № 7(4): С. 361–370. doi.org/10.15276/hait.07.2024.25. *Видання включено до переліку наукових фахових видань України, категорія «Б».*

<https://hait.od.ua/index.php/journal/article/view/232>

*Наукові праці, які засвідчують апробацію матеріалів дисертації:*

7. Shcherbakova, G., Antoshchuk, S., Koshutina, D., Sakhno, K., Kondratiev, S. (2025). Adaptive Clustering for Distribution Parameter Estimation in Technical Diagnostics. In: Dovgyi, S., Siemens, E., Globa, L., Kopyika, O., Stryzhak, O. (eds) *Applied Innovations in Information and Communication Technology. ICAIIT 2024. Lecture Notes in Networks and Systems*, vol 1338. Springer, Cham. doi.org/10.1007/978-3-031-89296-7\_19. *Видання проіндексовано у базі даних Scopus.*

[https://icaiit.org/proceedings/12th\\_ICAIIT\\_1/2\\_8\\_ICAIIT\\_2024\\_Part\\_2\\_paper\\_6.pdf](https://icaiit.org/proceedings/12th_ICAIIT_1/2_8_ICAIIT_2024_Part_2_paper_6.pdf)

8. Antoshchuk S., Shcherbakova G., Kondratyev S., Koshutina, D., Usov O. Wavelet transform based optimization method for Three-Dimensional computer vision // *Proceedings of the 12th International Conference Information Control Systems & Technologies (ICST 2024)*, Odesa, Ukraine, September 23–25, 2024, pp. 471–482. *Видання проіндексовано у базі даних Scopus.*

<https://ceur-ws.org/Vol-3790/paper41.pdf>

9. Kondratiev S., Koshutina D., Liubomska O., Baskov I. Research on Energy Efficiency of Wi-Fi IoT Systems on Renesas DA16200 Platform // *Proceedings of International Conference on Applied Innovation in IT*, Koethen, Germany, March 7, 2024, 12(1), pp. 143–148. *Видання проіндексовано у базі даних Scopus.*

[https://icaiit.org/proceedings/12th\\_ICAIIT\\_1/2\\_11\\_ICAIIT\\_2024\\_Part\\_2\\_paper\\_9.pdf](https://icaiit.org/proceedings/12th_ICAIIT_1/2_11_ICAIIT_2024_Part_2_paper_9.pdf)

10. Антощук С.Г., Щербакова Г.Ю., *Кондратьєв С.Б.*, Кошутіна Д.В. Наближений метод оцінки глибини об'єктів на основі вейвлет-перетворення. Матеріали X Міжнародної науково-практичної конференції Інформатика. Культура. Техніка. 2024, 25–27 вересня 2024 р. м. Одеса (Україна). 2024; 1(1): 197–201.

<https://ics60.aait.od.ua/zbirnik2024.pdf>

11. Світлана Антощук, Галина Щербакова, *Сергій Кондратьєв*, Дар'я Кошутіна, Олександр Усов. Метод оптимізації на основі вейвлет-перетворення для наближених методів оцінки глибини // Інформаційні управляючі системи та технології (ІУСТ-Одеса-2024): Матеріали XII Міжнародної науково-практичної конференції, 23–25 вересня 2024 р. м. Одеса (Україна). вип. ред. Вичужанін. – Одеса: Видавничий дім «Гельветика», 2024. – 334 с.123–127

<http://icst-conf.com/2024.pdf>

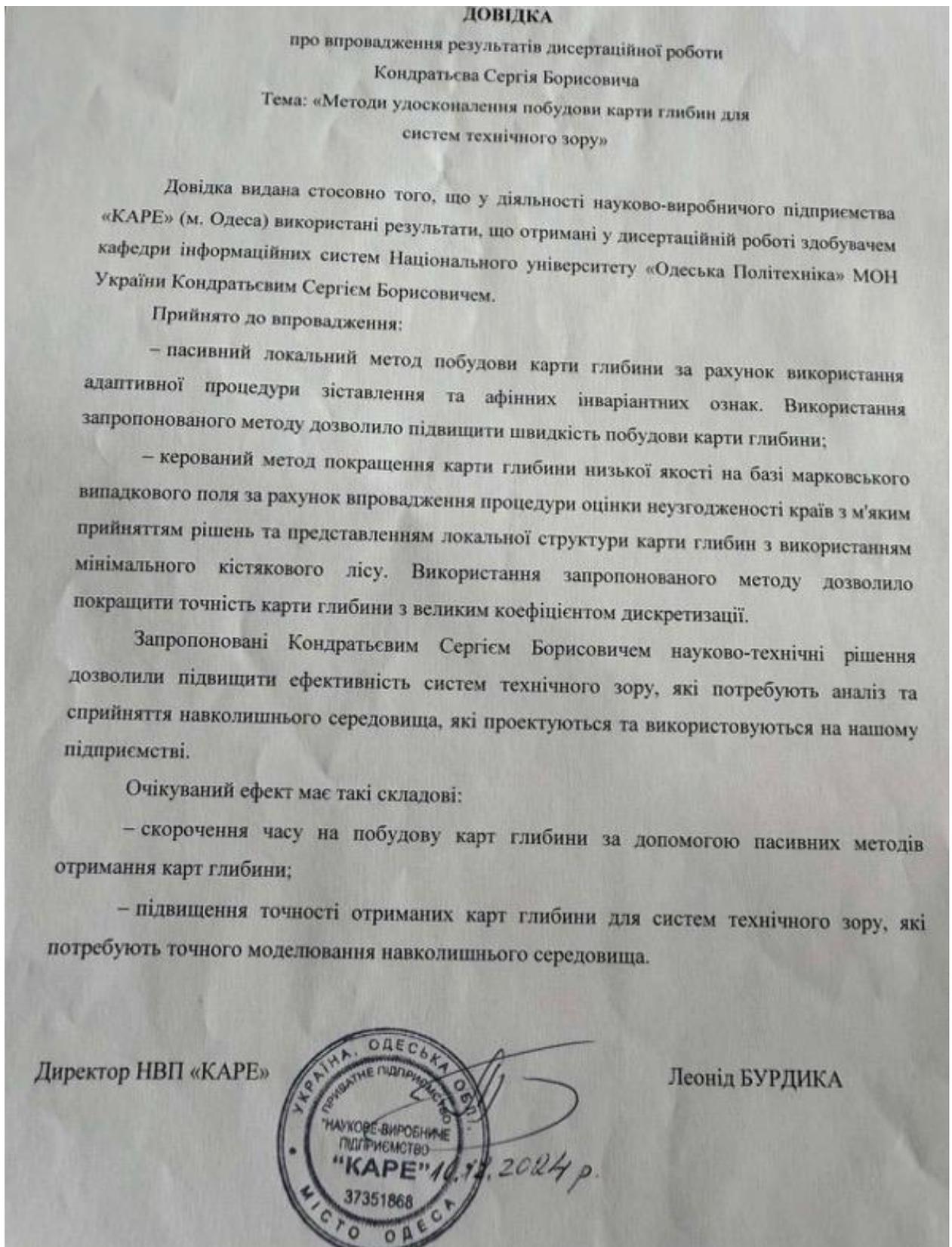
12. Антощук С.Г., Арсірій О.О., *Кондратьєв С.Б.* та ін. Гібридні моделі представлення та методи обробки зображень в просторі вейвлет-перетворення. 2020. 160 с.

<https://discovery.kpi.ua/Record/000634213>

13. Антощук С.Г., Арсірій О.О., *Кондратьєв С.Б.* та ін. Аналіз та розпізнавання зображень у просторі вейвлет-перетворення. 2022. 126 с.

<http://www.irbis-nbuv.gov.ua/publ/REF-0000823682>

**ДОДАТОК Б**  
**ДОВІДКА ПРО ВПРОВАДЖЕННЯ РЕЗУЛЬТАТІВ ДИСЕРТАЦІЙНОЇ**  
**РОБОТИ У ДІЯЛЬНІСТЬ НВП «КАРЕ»**





**ДОДАТОК Г**

**ДОВІДКА ПРО ВИКОРИСТАННЯ РЕЗУЛЬТАТІВ ДИСЕРТАЦІЙНОЇ  
РОБОТИ У НАУКОВО-ДОСЛІДНИЦЬКІЙ ДІЯЛЬНОСТІ  
НАЦІОНАЛЬНОГО УНІВЕРСИТЕТУ «ОДЕСЬКА ПОЛІТЕХНІКА»**

ДОВІДКА № 35

про використання результатів дисертаційної роботи  
Кондратьєва Сергія Борисовича  
«Методи удосконалення побудови карти глибин для систем технічного зору»  
у науково-дослідницькій діяльності  
Національного університету «Одеська політехніка»

Довідка видана стосовно того, що у науково-дослідницькій діяльності Національного університету «Одеська політехніка» використані наступні наукові результати, отримані у дисертаційній роботі Кондратьєва Сергія Борисовича:

– пасивний локальний метод побудови карти глибини на базі використання адаптивної процедури зіставлення та афінних інваріантних ознак та метод покращення часової узгодженості для послідовностей карт глибини на базі показника надійності глибини та використання динамічних атрибутів регіонів;

– керовані методи покращення карти глибини на базі марковського випадкового поля та процедури оцінки неузгодженості країв, а також представленням локальної структури карти глибин з використанням мінімального кістякового лісу.

Дисертацію виконано згідно тематичних планів НДР ОНПУ за період 2016 – 2025 рр. Кондратьєв Сергій Борисович був виконавцем у держбюджетних темах:

– № 719-61/161 – «ИНТЕЛЕКТУАЛЬНА ПІДТРИМКА ПРИЙНЯТТЯ РІШЕНЬ ПРИ ПРОЄКТУВАННІ КІСТКОВИХ ЗАМІННИКІВ ДЛЯ ЛІКУВАННЯ ВОЄННИХ ТРАВМ» (номер держ. реєстрації 0124U000388);

– № 716-145 «ІННОВАЦІЙНА ІНТЕЛЕКТУАЛЬНА ТЕХНОЛОГІЯ АНАЛІЗУ ТА СИНТЕЗУ ЕНЕРГОЕФЕКТИВНИХ ГІДРОАЕРОДИНАМІЧНИХ ЕЛЕМЕНТІВ ТА СИСТЕМ» (номер держ. реєстрації 0121U111790)

Проректор з наукової та наукової педагогічної роботи, д.т.н., професор



Дмитро ДМИТРИШИН

Керівник НДР  
№ 719-61/161 та 716-145  
д.т.н., професор

Світлана АНТОЩУК

**ДОДАТОК Д**  
**ДОВІДКА ПРО ВИКОРИСТАННЯ РЕЗУЛЬТАТІВ ДИСЕРТАЦІЙНОЇ**  
**РОБОТИ У НАВЧАЛЬНОМУ ПРОЦЕСІ НАЦІОНАЛЬНОГО**  
**УНІВЕРСИТЕТУ «ОДЕСЬКА ПОЛІТЕХНІКА»**

**ДОВІДКА**

про використання результатів дисертаційної роботи  
 Кондратьєва Сергія Борисовича  
 «Методи удосконалення побудови карти глибин для систем технічного зору»  
 у навчальному процесі  
 Національного університету «Одеська політехніка»

Довідка видана у результаті того, що в програмах, навчально-методичних матеріалах та курсах лекцій з дисциплін, що вивчаються на кафедрі Інформаційних систем Інституту комп'ютерних систем бакалаврами спеціальності 122 «Комп'ютерні науки» – «Інтелектуальний аналіз даних» та «Методи та системи штучного інтелекту» використовуються наукові результати, отримані у дисертації Кондратьєва Сергія Борисовича, а саме:

- дисципліна «Інтелектуальний аналіз даних»: базові алгоритми побудови карти глибин для систем технічного зору, тема «Методи та системи візуалізації даних»;
- дисципліна «Методи та системи штучного інтелекту»: методи удосконалення побудови карти глибин для систем технічного зору, тема «Інтелектуальні системи з технічним зором».

Використані результати дисертації Кондратьєва Сергія Борисовича «Методи удосконалення побудови карти глибин для систем технічного зору» свідчать про глибоке опрацювання літератури за тематикою дисертації. Висновки відзначаються науковою новизною, можливістю реалізації в інформаційних системах з необхідністю побудованої карти глибин, тому поліпшують матеріал означених дисциплін і сприяють підвищенню якості підготовки фахівців в галузі комп'ютерних наук та інформаційних технологій.

Перший проректор, проректор з науково-педагогічної та виховної роботи,  
 д.т.н., професор



Сергій НЕСТЕРЕНКО

Завідувачка кафедрою інформаційних систем ІКС,  
 д.т.н., професор

Олена АРСІРІЙ

04.12.2024 р.

## ДОДАТОК Е

### ФРАГМЕНТИ ВИХІДНОГО КОДУ ПРОГРАМНИХ ІНСТРУМЕНТАЛЬНИХ ЗАСОБІВ

```

import cv2

class StereoCamera:
    def __init__(self, left_source=0, right_source=1,
image_size=(640, 480), use_video_file=False, video_paths=None):
        self.use_video_file = use_video_file
        self.image_size = image_size

        if use_video_file and video_paths:
            self.left_cap = cv2.VideoCapture(video_paths["left"])
            self.right_cap = cv2.VideoCapture(video_paths["right"])
        else:
            self.left_cap = cv2.VideoCapture(left_source)
            self.right_cap = cv2.VideoCapture(right_source)
            self.left_cap.set(cv2.CAP_PROP_FRAME_WIDTH,
image_size[0])
            self.left_cap.set(cv2.CAP_PROP_FRAME_HEIGHT,
image_size[1])
            self.right_cap.set(cv2.CAP_PROP_FRAME_WIDTH,
image_size[0])
            self.right_cap.set(cv2.CAP_PROP_FRAME_HEIGHT,
image_size[1])

        def get_frame_pair(self):
            ret_left, left_frame = self.left_cap.read()
            ret_right, right_frame = self.right_cap.read()

            if not ret_left or not ret_right:
                raise RuntimeError("Не вдалося зчитати кадри з обох
джерел.")

            if self.image_size:
                left_frame = cv2.resize(left_frame, self.image_size)
                right_frame = cv2.resize(right_frame, self.image_size)

            return left_frame, right_frame

        def release(self):
            self.left_cap.release()

```

```

        self.right_cap.release()

    def is_opened(self):
        return self.left_cap.isOpened() and self.right_cap.isOpened()

if __name__ == "__main__":
    import time

    cam = StereoCamera(image_size=(640, 480))
    if not cam.is_opened():
        print("Камери не підключені!")
        exit(1)

    while True:
        try:
            left, right = cam.get_frame_pair()
            combined = cv2.hconcat([left, right])
            cv2.imshow("Stereo Capture", combined)

            key = cv2.waitKey(1)
            if key == 27:
                break
        except RuntimeError as e:
            print(str(e))
            break

    cam.release()
    cv2.destroyAllWindows()
import numpy as np
import cv2
import time

class ToFSensorSimulator:
    def __init__(self, depth_video_path=None, image_size=(640, 480),
                 fps=30):
        self.image_size = image_size
        self.fps = fps
        self.depth_video_path = depth_video_path
        self.cap = None
        if depth_video_path:
            self.cap = cv2.VideoCapture(depth_video_path)
            self.cap.set(cv2.CAP_PROP_FRAME_WIDTH, image_size[0])
            self.cap.set(cv2.CAP_PROP_FRAME_HEIGHT, image_size[1])

    def read_depth_frame(self):

```

```

    if not self.cap:
        return self._generate_synthetic_depth()
    ret, frame = self.cap.read()
    if not ret:
        self.cap.set(cv2.CAP_PROP_POS_FRAMES, 0)
        ret, frame = self.cap.read()
    gray = cv2.cvtColor(frame, cv2.COLOR_BGR2GRAY)
    depth = gray.astype(np.float32) / 255.0 * 5.0
    return depth

def _generate_synthetic_depth(self):
    t = time.time() % 10
    h, w = self.image_size
    xv, yv = np.meshgrid(np.linspace(-1, 1, w), np.linspace(-1,
1, h))
    depth = 2 + np.sin(2 * np.pi * (xv + yv + t / 10))
    depth = (depth - depth.min()) / (depth.max() - depth.min())
    depth = depth * 5.0
    return depth.astype(np.float32)

def release(self):
    if self.cap:
        self.cap.release()

if __name__ == "__main__":
    sensor = ToFSensorSimulator(depth_video_path=None)
    while True:
        depth = sensor.read_depth_frame()
        depth_vis = (depth / 5.0 * 255).astype(np.uint8)
        depth_colormap = cv2.applyColorMap(depth_vis,
cv2.COLORMAP_JET)
        cv2.imshow("ToF Depth Frame", depth_colormap)
        key = cv2.waitKey(int(1000 / sensor.fps))
        if key == 27:
            break
    sensor.release()
    cv2.destroyAllWindows()
# depth_map_toolkit/core/acquisition/video_loader.py

import cv2
import os
from typing import Optional, Tuple, List

class VideoLoader:

```

```

def __init__(self, left_path: str, right_path: Optional[str] =
None, image_size: Tuple[int, int] = (640, 480)):
    self.image_size = image_size

    if os.path.isdir(left_path):
        self.left_images = sorted([
            os.path.join(left_path, f) for f in
os.listdir(left_path)
            if f.lower().endswith(('.png', '.jpg', '.jpeg'))
        ])
        self.right_images = []
        if right_path and os.path.isdir(right_path):
            self.right_images = sorted([
                os.path.join(right_path, f) for f in
os.listdir(right_path)
                if f.lower().endswith(('.png', '.jpg', '.jpeg'))
            ])
        self.frame_mode = True
        self.index = 0
    else:
        self.left_cap = cv2.VideoCapture(left_path)
        self.right_cap = cv2.VideoCapture(right_path) if
right_path else None
        self.left_cap.set(cv2.CAP_PROP_FRAME_WIDTH,
image_size[0])
        self.left_cap.set(cv2.CAP_PROP_FRAME_HEIGHT,
image_size[1])
        if self.right_cap:
            self.right_cap.set(cv2.CAP_PROP_FRAME_WIDTH,
image_size[0])
            self.right_cap.set(cv2.CAP_PROP_FRAME_HEIGHT,
image_size[1])
        self.frame_mode = False

    def get_next_frame(self) -> Tuple:
        if self.frame_mode:
            if self.index >= len(self.left_images):
                raise StopIteration("Досягнуто кінця послідовності
зображень.")

            left = cv2.imread(self.left_images[self.index])
            right = None
            if self.right_images:
                right = cv2.imread(self.right_images[self.index])
            self.index += 1
        else:

```

```

        ret_left, left = self.left_cap.read()
        if not ret_left:
            raise StopIteration("Відеопотік лівої камери
завершено.")
        right = None
        if self.right_cap:
            ret_right, right = self.right_cap.read()
            if not ret_right:
                raise StopIteration("Відеопотік правої камери
завершено.")

        if self.image_size:
            left = cv2.resize(left, self.image_size)
            if right is not None:
                right = cv2.resize(right, self.image_size)

        return (left, right)

def reset(self):
    if self.frame_mode:
        self.index = 0
    else:
        self.left_cap.set(cv2.CAP_PROP_POS_FRAMES, 0)
        if self.right_cap:
            self.right_cap.set(cv2.CAP_PROP_POS_FRAMES, 0)

def release(self):
    if not self.frame_mode:
        self.left_cap.release()
        if self.right_cap:
            self.right_cap.release()
# depth_map_toolkit/core/preprocessing/wavelet.py

import cv2
import numpy as np
import pywt

class WaveletPreprocessor:
    def __init__(self, wavelet: str = 'haar', level: int = 1, mode:
str = 'symmetric'):
        self.wavelet = wavelet
        self.level = level
        self.mode = mode

    def decompose(self, image: np.ndarray) -> np.ndarray:

```

```

"""
Виконує вейвлет-декомпозицію зображення.
Повертає лише детальні компоненти (горизонтальні,
вертикальні, діагональні).
"""
if image.ndim == 3:
    image = cv2.cvtColor(image, cv2.COLOR_BGR2GRAY)
    image = image.astype(np.float32) / 255.0
    coeffs = pywt.wavedec2(image, wavelet=self.wavelet,
level=self.level, mode=self.mode)
    _, details = coeffs[0], coeffs[1:]
    detail_maps = []
    for cH, cV, cD in details:
        detail = np.sqrt(cH**2 + cV**2 + cD**2)
        detail_maps.append(detail)
    combined = np.mean(detail_maps, axis=0)
    combined = cv2.resize(combined, image.shape[:-1])
    return (combined * 255).astype(np.uint8)

def extract_edges(self, image: np.ndarray, threshold: float =
30.0) -> np.ndarray:
    """
    Повертає бінарну карту контурів на основі вейвлет-компонент.
    """
    detail_map = self.decompose(image)
    _, binary_edges = cv2.threshold(detail_map, threshold, 255,
cv2.THRESH_BINARY)
    return binary_edges

def enhance_texture_regions(self, image: np.ndarray) ->
np.ndarray:
    """
    Використовує деталі вейвлетів для підсилення текстурованих
ділянок.
    """
    details = self.decompose(image).astype(np.float32)
    original = cv2.cvtColor(image,
cv2.COLOR_BGR2GRAY).astype(np.float32)
    enhanced = cv2.normalize(original + 0.5 * details, None, 0,
255, cv2.NORM_MINMAX)
    return enhanced.astype(np.uint8)

if __name__ == "__main__":
    img = cv2.imread("sample.png")
    wp = WaveletPreprocessor(level=2)

```

```

details = wp.decompose(img)
edges = wp.extract_edges(img)
enhanced = wp.enhance_texture_regions(img)

cv2.imshow("Original", img)
cv2.imshow("Wavelet Details", details)
cv2.imshow("Wavelet Edges", edges)
cv2.imshow("Enhanced Texture", enhanced)
cv2.waitKey(0)
cv2.destroyAllWindows()
# depth_map_toolkit/core/preprocessing/filtering.py

import cv2
import numpy as np

def bilateral_filter_depth(depth_map: np.ndarray, diameter: int = 9,
                           sigma_color: float = 75, sigma_space:
float = 75) -> np.ndarray:
    """
    Білатеральне згладжування карти глибини.
    :param depth_map: одноканальна карта глибини (float32 або uint8)
    :return: згладжена карта глибини
    """
    if depth_map.dtype != np.uint8:
        norm = cv2.normalize(depth_map, None, 0, 255,
cv2.NORM_MINMAX)
        depth_map = norm.astype(np.uint8)
    filtered = cv2.bilateralFilter(depth_map, diameter, sigma_color,
sigma_space)
    return filtered.astype(np.float32)

def guided_filter_depth(depth_map: np.ndarray, guidance: np.ndarray,
                        radius: int = 5, eps: float = 1e-2) ->
np.ndarray:
    """
    Керована фільтрація зображення глибини за допомогою керівного
зображення.
    :param depth_map: float32
    :param guidance: uint8 or float32 (часто - RGB або grayscale)
    :return: згладжена карта глибини
    """
    import cv2.ximgproc as xip
    if guidance.ndim == 3:

```

```

        guidance = cv2.cvtColor(guidance, cv2.COLOR_BGR2GRAY)
        guidance = guidance.astype(np.float32) / 255.0
        return xip.guidedFilter(guidance, depth_map, radius, eps)

def mask_weighted_filter(depth_map: np.ndarray, reliability_mask:
np.ndarray,
                        kernel_size: int = 5) -> np.ndarray:
    """
    Фільтрація карти глибини з урахуванням маски довіри.
    Значення з низькою довірою пригнічуються, як у дисертації.
    :param depth_map: float32
    :param reliability_mask: 0 або 1 (uint8)
    :return: уточнена карта глибини
    """
    kernel = np.ones((kernel_size, kernel_size), dtype=np.float32)
    reliable = depth_map * reliability_mask
    weight_sum = cv2.filter2D(reliability_mask.astype(np.float32), -
1, kernel)
    value_sum = cv2.filter2D(reliable.astype(np.float32), -1,
kernel)
    result = np.zeros_like(depth_map)
    with np.errstate(divide='ignore', invalid='ignore'):
        result = np.where(weight_sum > 0, value_sum / weight_sum,
depth_map)
    return result

if __name__ == "__main__":
    dmap = cv2.imread("depth.png",
cv2.IMREAD_GRAYSCALE).astype(np.float32)
    reliability = (dmap > 0).astype(np.uint8)

    bilateral = bilateral_filter_depth(dmap)
    masked = mask_weighted_filter(dmap, reliability)

    cv2.imshow("Original", dmap.astype(np.uint8))
    cv2.imshow("Bilateral", bilateral.astype(np.uint8))
    cv2.imshow("Mask Weighted", masked.astype(np.uint8))
    cv2.waitKey(0)
    cv2.destroyAllWindows()
# depth_map_toolkit/core/features/affine_features.py

import cv2
import numpy as np
from typing import Tuple, List

```

```

class AffineFeatureExtractor:
    def __init__(self, method: str = 'SIFT', max_features: int =
500):
        if method.upper() == 'SIFT':
            self.detector = cv2.SIFT_create(nfeatures=max_features)
        elif method.upper() == 'ORB':
            self.detector = cv2.ORB_create(nfeatures=max_features)
        else:
            raise ValueError("Підтримуються лише SIFT або ORB")
        self.method = method.upper()

    def detect_and_compute(self, image: np.ndarray) ->
Tuple[List[cv2.KeyPoint], np.ndarray]:
        if image.ndim == 3:
            image = cv2.cvtColor(image, cv2.COLOR_BGR2GRAY)
        keypoints, descriptors =
self.detector.detectAndCompute(image, None)
        return keypoints, descriptors

    def match_features(self, desc1: np.ndarray, desc2: np.ndarray,
ratio_thresh: float = 0.75) -> List[cv2.DMatch]:
        if self.method == 'SIFT':
            matcher = cv2.BFMatcher(cv2.NORM_L2)
        elif self.method == 'ORB':
            matcher = cv2.BFMatcher(cv2.NORM_HAMMING)
        matches = matcher.knnMatch(desc1, desc2, k=2)
        good_matches = []
        for m, n in matches:
            if m.distance < ratio_thresh * n.distance:
                good_matches.append(m)
        return good_matches

    def draw_matches(self, img1: np.ndarray, kp1, img2: np.ndarray,
kp2, matches: List[cv2.DMatch]) -> np.ndarray:
        return cv2.drawMatches(img1, kp1, img2, kp2, matches, None,
flags=cv2.DrawMatchesFlags_NOT_DRAW_SINGLE_POINTS)

if __name__ == "__main__":
    img1 = cv2.imread("left.png")
    img2 = cv2.imread("right.png")

    feature_extractor = AffineFeatureExtractor(method='SIFT')

```

```

kp1, des1 = feature_extractor.detect_and_compute(img1)
kp2, des2 = feature_extractor.detect_and_compute(img2)
matches = feature_extractor.match_features(des1, des2)

vis = feature_extractor.draw_matches(img1, kp1, img2, kp2,
matches)
print(f"Знайдено пар ознак: {len(matches)}")
cv2.imshow("Feature Matches", vis)
cv2.waitKey(0)
cv2.destroyAllWindows()
# depth_map_toolkit/core/features/disparity_cost.py

import numpy as np
import cv2

class DisparityCostComputer:
    def __init__(self, max_disparity: int = 64, block_size: int = 5,
method: str = "SSD"):
        """
        :param max_disparity: Максимальний діапазон зміщення (у
пікселях)
        :param block_size: Розмір вікна (має бути непарним)
        :param method: Метод обчислення витрат ('SSD', 'SAD', 'NCC')
        """
        self.max_disp = max_disparity
        self.block_size = block_size
        self.method = method.upper()

    def compute_cost_volume(self, left_img: np.ndarray, right_img:
np.ndarray) -> np.ndarray:
        if left_img.ndim == 3:
            left_img = cv2.cvtColor(left_img, cv2.COLOR_BGR2GRAY)
        if right_img.ndim == 3:
            right_img = cv2.cvtColor(right_img, cv2.COLOR_BGR2GRAY)

        h, w = left_img.shape
        half_block = self.block_size // 2
        cost_volume = np.zeros((h, w, self.max_disp),
dtype=np.float32)

        left_padded = cv2.copyMakeBorder(left_img, half_block,
half_block, half_block + self.max_disp, half_block,
cv2.BORDER_REFLECT)

```

```

        right_padded = cv2.copyMakeBorder(right_img, half_block,
        half_block, half_block + self.max_disp, half_block,
        cv2.BORDER_REFLECT)

        for d in range(self.max_disp):
            left_crop = left_padded[:, self.max_disp: self.max_disp +
w]
            right_crop = right_padded[:, self.max_disp - d:
self.max_disp - d + w]
            if self.method == "SSD":
                cost = cv2.boxFilter((left_crop - right_crop) ** 2, -
1, (self.block_size, self.block_size))
            elif self.method == "SAD":
                cost = cv2.boxFilter(np.abs(left_crop - right_crop),
-1, (self.block_size, self.block_size))
            elif self.method == "NCC":
                cost = self._ncc_cost(left_crop, right_crop)
            else:
                raise ValueError("Метод повинен бути один з: SSD, SAD,
NCC")

            cost_volume[:, :, d] = cost
        return cost_volume

    def _ncc_cost(self, img1, img2):
        img1 = img1.astype(np.float32)
        img2 = img2.astype(np.float32)
        mean1 = cv2.boxFilter(img1, -1, (self.block_size,
self.block_size))
        mean2 = cv2.boxFilter(img2, -1, (self.block_size,
self.block_size))
        std1 = np.sqrt(cv2.boxFilter((img1 - mean1) ** 2, -1,
(self.block_size, self.block_size)))
        std2 = np.sqrt(cv2.boxFilter((img2 - mean2) ** 2, -1,
(self.block_size, self.block_size)))
        numerator = cv2.boxFilter((img1 - mean1) * (img2 - mean2), -
1, (self.block_size, self.block_size))
        denominator = std1 * std2 + 1e-5
        ncc = 1.0 - numerator / denominator
        return ncc

if __name__ == "__main__":
    left = cv2.imread("left.png")
    right = cv2.imread("right.png")
    cost_computer = DisparityCostComputer(max_disparity=64,
block_size=7, method="SSD")

```

```

cost_volume = cost_computer.compute_cost_volume(left, right)
cost_map = cost_volume[:, :, 10]
norm_cost = cv2.normalize(cost_map, None, 0, 255,
cv2.NORM_MINMAX).astype(np.uint8)
cv2.imshow("Cost Map at Disparity=10", norm_cost)
cv2.waitKey(0)
cv2.destroyAllWindows()
# depth_map_toolkit/core/features/edge_consistency.py

import cv2
import numpy as np

def compute_edge_map(image: np.ndarray, blur: bool = True) ->
np.ndarray:
    """
    Обчислює карту контурів зображення за допомогою оператора Собеля.
    :param image: кольорове або сіре зображення
    :return: карта контурів (float32)
    """
    if image.ndim == 3:
        image = cv2.cvtColor(image, cv2.COLOR_BGR2GRAY)

    if blur:
        image = cv2.GaussianBlur(image, (3, 3), 0)

    sobelx = cv2.Sobel(image, cv2.CV_32F, 1, 0, ksize=3)
    sobely = cv2.Sobel(image, cv2.CV_32F, 0, 1, ksize=3)
    gradient_magnitude = cv2.magnitude(sobelx, sobely)
    return gradient_magnitude

def compute_edge_inconsistency(rgb_image: np.ndarray, depth_map:
np.ndarray, threshold: float = 0.1) -> np.ndarray:
    """
    Обчислює карту неузгодженості країв між RGB-зображенням і картою
    глибини.
    :param rgb_image: вхідне кольорове зображення
    :param depth_map: float32 карта глибини (в нормалізованому або
    абсолютному масштабі)
    :param threshold: поріг на відхилення градієнтів
    :return: бінарна карта неузгоджених країв (uint8)
    """
    edge_rgb = compute_edge_map(rgb_image)
    edge_depth = compute_edge_map(depth_map)

```

```

    edge_rgb = cv2.normalize(edge_rgb, None, 0, 1, cv2.NORM_MINMAX)
    edge_depth = cv2.normalize(edge_depth, None, 0, 1,
cv2.NORM_MINMAX)

    diff = np.abs(edge_rgb - edge_depth)
    inconsistency = (diff > threshold).astype(np.uint8) * 255
    return inconsistency

def soft_edge_difference(rgb_image: np.ndarray, depth_map:
np.ndarray) -> np.ndarray:
    """
    Повертає кількісну карту неузгодженості між RGB та глибинними
краями (float32).
    :param rgb_image: кольорове зображення
    :param depth_map: float32 карта глибини
    :return: карта відносної відмінності (0...1)
    """
    edge_rgb = compute_edge_map(rgb_image)
    edge_depth = compute_edge_map(depth_map)

    edge_rgb = cv2.normalize(edge_rgb, None, 0, 1, cv2.NORM_MINMAX)
    edge_depth = cv2.normalize(edge_depth, None, 0, 1,
cv2.NORM_MINMAX)

    diff = np.abs(edge_rgb - edge_depth)
    return diff.astype(np.float32)

if __name__ == "__main__":
    rgb = cv2.imread("color.png")

```