



KAPITEL 5 / CHAPTER 5⁵
**METHODS FOR FINDING KEY POINTS IN AMBER SAMPLE IMAGES AS
A BASIS FOR THREE-DIMENSIONAL MODELING**

DOI: 10.30890/2709-2313.2025-43-01-020

Бурштин – це скам'яніла викопна смола хвойних дерев, яка утворилась 45-50 млн. років тому. Також поширена інша назва – сукциніт, що з латині означає сосни. Свою дорогоцінність отримав через те, що на його формування пішли мільйони років та обмежені за площею ареали поширення родовищ бурштину. Також на загальну ціну впливає колір зразку та вміст інклюзій у вигляді тогочасних комах чи малих тварин (ящірки, жаби), які навіки застигли у бурштиновій смолі мільйони років тому.

Видобування бурштину доволі непростий процес, так як він залягає глибоко в землі та в обмеженій концентрації. Та сама експлуатація родовищ потребує спеціалізованого обладнання. Механізоване розорювання ґрунтів, вирубування лісів та промивання порід великою кількістю води, призводить до руйнування цілих екосистем, посилює ерозійні процеси та призводить до деградації землі. Відбувається порушення гідрологічного режиму ареалу, змінюється рівень ґрунтових вод, що веде до негативного впливу на усі живі організми.

Застосувань бурштину купа, він широко використовується в ювелірному мистецтві, народній медицині, косметології, парфумерії, а також у електроніці та палеонтологічних дослідженнях. Перед практичним застосуванням, наприклад в ювелірній справі, його як правило обробляють майстри та надаючи потрібну форму. Під час такої обробки втрачається значна частина маси, тому особливої актуальності набуває розроблення та покращення технологій оцінювання форми та візерунку зразків. Використання таких технологій дасть змогу раціонально витратити сировину, що в свою чергу дозволить задовольнити більший попит за рахунок її менших обсягів. Зі зменшенням видобувних обсягів зменшиться і навантаження на екологію, так як саме видобування завдає непоправної шкоди

⁵*Authors: Ryzhuk Andrii Viktorovych*



навколишньому середовищу.

Для більш точної оцінки каменів використовують наступні методи:

- Комп'ютерна томографія. Такий метод дозволяє створювати тривимірні моделі зразка бурштину, виявляти та досліджувати внутрішню структуру каменю. Недоліком комп'ютерної томографії є висока вартість спеціалізованого обладнання та вузькоспеціалізоване програмне забезпечення.

- Інфрачервона спектроскопія. Метод дослідження дозволяє виявляти не тільки структурні особливості каменю, а й хімічний склад зразку. Недоліком методу є складність в інтерпретації спектрів та потреба у висококваліфікованих кадрах.

- Оптична мікроскопія. Такий метод дозволяє детально досліджувати поверхню зразку до найменших структур. До недоліків належить обмеженість глибини аналізу.

- Стереобачення. Метод що базується на використанні стереокамери, що сканує об'єкт та відтворює детальну тривимірну модель. До недоліків варто віднести високу вартість спеціалізованого обладнання.

З огляду на обмеженість сучасних підходів постає задача у створенні ефективного та доступного автоматизованого рішення для дослідження зразків бурштину на основі сучасних технологій технічного зору та аналізу їх тривимірної моделі. За допомогою тривимірних моделей досить легко оцінювати форми, вкраплення, об'єм та колір зразку без безпосередньої взаємодії зі зразком. Така інформація надає промислову оцінку про родовище та дає змогу розрахувати майбутній виріб з мінімальними витратами матеріалу, який буде втрачений при обробці.

Метою проведених досліджень є оцінювання точнісних та часових характеристик найпоширеніших методів виявлення ключових точок на парі зображень, також їх удосконалення в напрямі мінімізації репроекційної помилки та використання цих методів в задачах тривимірної візуалізації зразків бурштину.

Для досягнення поставленої мети необхідно розв'язати наступні завдання:



1. Отримання секвенції зображень досліджуваного зразка бурштину;
2. Дослідження та порівняння двох методів визначення ключових точок SIFT та ORB, обґрунтування вибору оптимального методу на основі аналізу точнісних та часових характеристик;
3. Поліпшення обраного методу в контексті точності побудови розрідженою хмари точок та мінімізації репроекційної помилки;
4. Відновлення просторового положення ключових точок методом їх триангуляції.

Методологія

Математичний опис технічних засобів формувань зображень ґрунтується на моделі камери-обскури. Проектування тривимірних об'єктів, а саме точок, які ним належать на площину можна представити таким рівнянням [1]:

$$s p = A[R|t]R_w, \quad (1)$$

де R_w - просторова точка, що задана в глобальній системі координат, p – точка, яка є проекцією на двомірній площині, s – параметр масштабування, A – внутрішні параметри камери(матриця 3×3), $R|t$ – зовнішні параметри камери(обертання та переміщення). Перед проведенням тривимірної реконструкції потрібно відкалібрувати камеру. Для отримання параметрів внутрішньої матриці A використовується тестове зображення – на якому зображено патерн з об'єктів однакового розміру. Найпопулярнішим є патерн для калібрування – шахова сітка. Важливим правилом для калібрування тестовим зображенням є відсутність однозначності при повороті на 180 градусів. Далі відбувається фотофіксація цього шаблону з різних ракурсів та визначаються внутрішні параметри камери [1].

Після калібрування відбувається безпосередня фотозйомка зразка бурштину та за набором таких зображень працює основний метод тривимірної реконструкції сцени SFM (structure from motion). Основна ідея якого полягає у відновленні геометрії сцени та визначення параметрів руху та параметрів камери. На першому етапі визначаються ключові точки одним з детекторів та обчислюються їх дескриптори. Після чого на основі цих дескрипторів



відбувається пошук відповідності між ключовими точками на парах зображень та знаходження есенціальної матриці. На основі відповідних точок визначаються відносні параметри руху камери, а саме обертання та переміщення, та дозволяє побудувати початкову розріджену хмару точок що відображає правильність побудови сканованої сцени. Далі кожне наступне зображення за таким самим алгоритмом додаються до сцени доповнюючи новими тривимірними точками. При кожному додаванні до сцени нових тривимірних точок виконується корегування пучка, тобто спільна нелінійна оптимізація, що мінімізує похибку шляхом одночасного уточнення параметрів камер й координат просторових точок. Цей параметр відображає точнісні характеристики процесу триангуляції для пари двомірних зображень[2]. Для того щоб мінімізувати репроекційну помилку потрібно забезпечити високоточне калібрування камери, коректне та точне зіставлення контрольних точок на парі зображень. Для знаходження та зіставлення використовується один з відомих алгоритмів: SIFT(scale-invariant feature transform) та ORB(oriented FAST and Rotated BRIEF).

Метод SIFT використовується для знаходження ключових точок на зображенні та їх дескрипторів. Такий метод дозволяє виявляти найстійкіші точки які є стійкими до обертання, масштабування та інших афінних перетворень. Пошук ключових точок відбувається за допомогою визначення локальних екстремумів. До зображення застосовується Гауссів шум та за допомогою «Різниці Гауса», або DoG визначаються найстійкіші точки. Далі для кожної зі знайдених точок розраховується гістограма напрямку базуючись на магнітуді та орієнтації:

$$Magnitude = \sqrt{G_x^2 + G_y^2}, \quad (2)$$

$$Orientation = \arctan\left(\frac{G_y}{G_x}\right), \quad (3)$$

де значення G_x та G_y - часткові похідні функції інтенсивності за двома просторовими координатами x та y . Орієнтацією є напрямок градієнт функції інтенсивності, а величиною є значення його магнітуди. Отримавши такі значення створюється дескриптор для визначеної точки. В класичному варіанті SIFT



дескриптори створюють масив з 16 локальних гістограм у вигляді сітки 4x4, де в кожному осередку накопичуються значення за 8 напрямками градієнтів орієнтації. На основі цих гістограм відбувається знаходження та зіставлення ключових точок на парі зображень[3].

Іншим не менш популярним метод для виявлення ключових точок є ORB. Він поєднує в собі 2 методи: FAST(features from accelerated segment text) для знаходження точок та BRIEF(binary robust independent elementary features) для обрахунку їх дескрипторів. У методу FAST ключовими точками є кути. Ці кути визначаються за допомогою побудови навколо вибраного пікселя кола Брезенмеха. Після чого порівнюються інтенсивності пікселів кола з центральним пікселем. Логіка вибору наступна: якщо вісім пікселів темніші чи світліші то метод позначає точку як ключову. Метод ORB використовує модифікований FAST. В модифікованому алгоритм вибору ключових точок базується на побудові тих самих зображень зі зменшенням роздільної здатності. На кожному етапі поділу відбувається пошук точок за допомогою звичайного FAST. Завдяки такій особливості цей метод є частково інваріантним до масштабування. Для знайдених точок обраховуються дескриптори за допомогою BRIEF, а саме визначається напрямок градієнта зміни інтенсивності.

$$m_{pq} = \sum_{x,y} x^p y^q I(x, y), \quad (4)$$

Та на основі моментів яскравості зображення визначається центроїд:

$$C = \left(\frac{m_{10}}{m_{00}}, \frac{m_{01}}{m_{00}} \right), \quad (5)$$

а також визначається напрям вектора від центра області до його центра мас:

$$\theta = \text{atan} 2(m_{01}, m_{10}). \quad (6)$$

Після обрахунку параметрів формується бінарний вектор ознак за правилом, якщо інтенсивність пікселю x менша від пікселю y то у вектор записується 1 в іншому випадку 0. Результатом такої обробки стає дескриптор розміром 256 біт.



Такий вектор ознак легко порівнювати та знаходити відповідності за допомогою відстані Хемінга[4].

Для покращення зіставлення та перевірки надійності відповідності точок на парі зображень також рекомендовано вести додаткову перевірку з пороговим значення близько 0.8. Це усуває 90% помилок з усієї вибірки зіставлених результатів. Далі на основі ключових точок, внутрішніх параметрів камери обраховується есенціальна матриця, коректне зіставлення точок на парі зображень грає ключову роль в її розрахунку. При помилковому визначенні есенціальна матриця буде неточною, або взагалі не правильною через що буде порушено реальне положення камер і це призведе до великих значень репроекційної помилки і ставитиме під сумнів весь результат тріангуляції зразка бурштину. Тому важливо забезпечити дійсно правильні ключову точки які будуть використовуватись для обрахунку есенціальної матриці.

Проведено порівняння двох найпопулярніших методів для виявлення ключових ознак ORB та SIFT, останній з яких був доповнений додатковою фільтрацією для більш точного зіставлення ключових точок. Тому першим етапом було виконано калібрування камери за допомогою дощечки з шаховим порядком. Розмір дощечки 9x7 квадратів, що відповідає критерію неоднозначності при повороті. Було проведено фотофіксацію з різних ракурсів калібрувальних дощечок та проведено калібрування камери Nikon D3100 kit. За допомогою загального алгоритму для калібрування камери та функції `cv2.calibrateCamera` в Python було отримано значення репроекційної помилки 0.3153px, що є прийнятним в аспекті точності, а також внутрішні параметри камери. Після проведення калібрування відбувається наповнення секвенції зображень для зразка бурштину під різними ракурсами та до даної групи зображень застосовуються методи SIFT та ORB для отримання точніших та часових результатів.

Таблиця 1 - Часові та кількісні характеристики пошуку ключових точок за допомогою метода SIFT

Номер пари зображень для пошуку ключових точок (ітерація пошуку)	Час пошуку ключових точок на зображеннях,с		Кількість знайдених ключових точок		Кількість співпадінь ключових точок для двох зображень
	Перше зображення пари	Друге зображення пари	Перше зображення пари	Друге зображення пари	
0	1,61	1,64	5965	7636	122
1	1,56	1,55	7636	10588	61
2	1,62	1,55	10588	3063	46
3	1,57	1,62	3063	4403	145
4	1,60	1,60	4403	5458	144
5	1,54	1,56	5458	5091	67
6	1,58	1,62	5091	19005	43
7	1,63	1,56	19005	9489	286
8	1,58	1,56	9489	3676	293
9	1,54	1,53	3676	2559	84
10	1,57	1,54	2559	2677	36
11	1,54	1,53	2677	3402	37
12	1,53	1,53	3402	1796	51
13	1,53	1,59	1796	3491	28
14	1,61	1,55	3491	1862	39
15	1,53	1,56	1862	5965	36
Загальна	25,22	25,17			1518

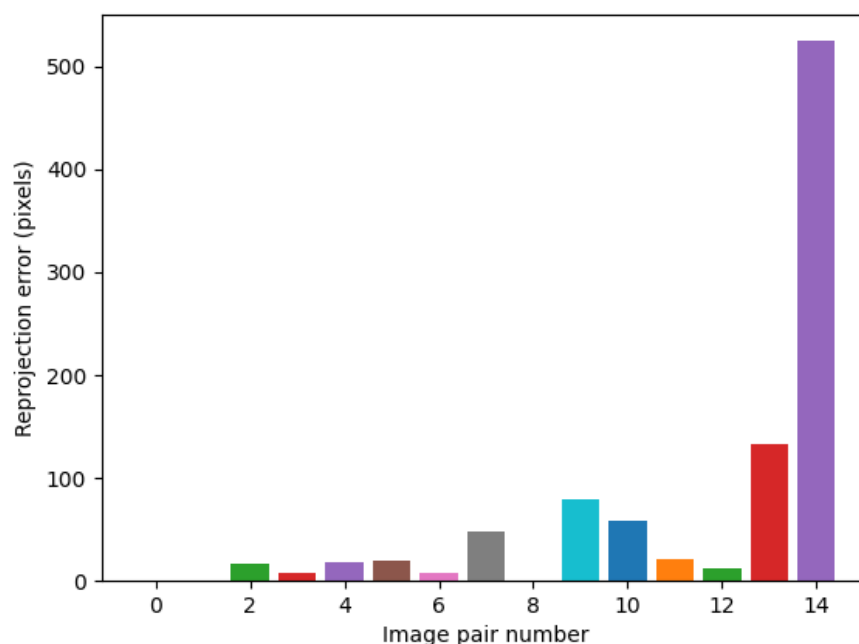


Рисунок 1 – Репроекційні похибки SFM алгоритму з використанням SIFT.



Таблиця 2 - Часові та кількісні характеристики пошуку ключових точок за допомогою метода ORB

Номер пари зображень для пошуку ключових точок (ітерація пошуку)	Час пошуку ключових точок на зображеннях,с		Кількість знайдених ключових точок		Кількість співпадінь ключових точок для двох зображень
	Перше зображення пари	Друге зображення пари	Перше зображення пари	Друге зображення пари	
0	0,276781082	0,101001501	17384	19411	159
1	0,092030048	0,092998981	19411	19339	27
2	0,097999334	0,075999022	19339	9452	32
3	0,069000483	0,081983328	9452	13500	172
4	0,074025869	0,086037397	13500	16590	171
5	0,082998991	0,080023766	16590	14165	42
6	0,088000059	0,111000299	14165	19929	27
7	0,103029251	0,094996691	19929	18910	310
8	0,087000847	0,078025818	18910	10568	325
9	0,074026346	0,073998451	10568	6697	52
10	0,066026449	0,070000172	6697	6431	25
11	0,066026688	0,074973106	6431	8144	37
12	0,068027735	0,068000555	8144	4853	43
13	0,062025309	0,078000307	4853	10039	14
14	0,076025486	0,067998886	10039	4526	15
15	0,067026615	0,08699894	4526	17384	29
Загальна	1,450050592	1,32203722			1480

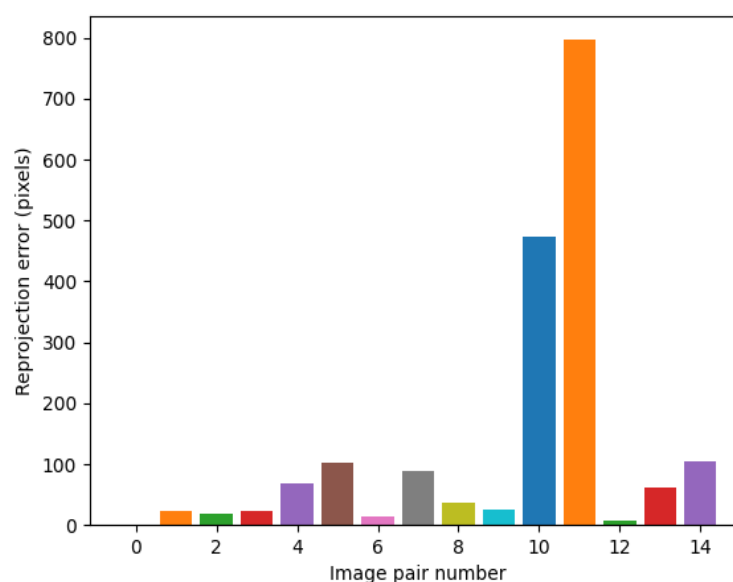


Рисунок 2 – Репроекційні похибки SFM алгоритму з використанням ORB.



При порівнянні двох методів SIFT та ORB за параметрами швидкодії та точності, а також значення репроекційної похибки було визначено наступне:

1. Відповідно до даних про швидкодію обчислення ключових точок від номера ітерації, метод ORB демонструє менші часові затрати на їх пошук в порівнянні SIFT (таб.1, таб.2)

2. Метод ORB також демонструє більшу кількість знайдених контрольних точок, в порівнянні з SIFT (таб.1, таб.2)

3. Більша загальна кількість коректних відповідностей протягом усього дослідження залишається у метода SIFT, так як він більш адаптований до інваріантних пар зображень ніж ORB і має більш кращі алгоритми виявлення стійких ключових точок на основі екстремумів (таб.1, таб.2).

В даному завданні реконструкції на основі SFM алгоритму важливим критерієм є саме точність, а не кількість чи швидкодія, тому порівнюючи графіки рис.1 та рис.2 можна побачити, що більше викидів відбувається при використанні метода ORB та в загальному по всім значенням репроекційної похибки SIFT демонструє дещо кращі результати. Проте все одно більшість результатів знаходяться в діапазоні 0.87рх до 100рх. Такі великі значення репроекційної похибки свідчать про те, що помилково відбулись хибні зіставлення ключових точок між парами зображень та есенціальна матриця була хибно порахована.

На хибні зіставлення, саме в зразках бурштину можуть впливати особливості самої поверхні зразка. Різноманітні візерунки у вигляді текстури, тріщин, дрібних включень можуть мати подібність на деяких ділянках, що в результаті створює дуже схожі дескриптори і це може заплутати алгоритм зіставлення та помилково визначити хибну пару. Також на полірованих поверхнях можуть з'являтися відблиски та бліки. В загальному патерн білків виглядає однаково для усіх випадків. Це світла пляма круглої або овальної форми на темному фоні. Такі артефакти можуть виникати на різних ділянках і мати схожі дескриптори. Що при фільтруванні коефіцієнтом відношення 0.8 ніяк не усунеться і відбудеться хибний збіг, який потім буде використаний при



прорахунку есенціальної матриці та вплине на значення репроекційної похибки. Тому важливо забезпечити більш стійку фільтрацію таких хибних збігів, що не видаляються коефіцієнтом відношення 0.8.

Для зменшення репроекційної помилки потрібно шукати ключові точки не просто в одному напрямку, а в обох. Відбувається симетрична перевірка відповідностей між точками p_i на основі дескрипторів (на першому зображенні) та q_i (на другому). Точка вважається коректною якщо p_i є найкращим збігом для q_i в одному напрямку та q_i є найкращим збігом для p_i в іншому напрямку. Таким чином після перевірки отримуються тільки взаємно узгоджені й локально унікальні відповідності, що істотно знижує кількість неправильних збігів зменшуючи помилкові розрахунки та вплив на непроекційну похибку.

Іншою перевіркою, що дозволить підвищити кількість коректних збігів, це перевірка їх сусідів. Для кожного збігу знаходиться k кількість найближчих сусідів на першому зображенні та на другому зображенні знаходяться та ж сама k кількість відповідних сусідів, що і на першому. Далі для кожного сусіда обраховується кут напрямку та відстань. Потім відбувається пошук локальної геометрії навколо збігу в значенні кута та масштабу та визначається наскільки змінилось. Тобто якщо збіг дійсно коректний, то сусіди на обох зображеннях матимуть приблизно однакові локальні геометричні властивості, проте якщо збіг хибний відповідні сусіди будуть в абсолютно інших місцях та демонструватимуть великі кутові та масштабні значення. На основі цих властивостей вдасться ще більш точно визначити коректні збіги між парою зображень та відсіяти помилкові, які не були відсіяні на етапі двостороннього пошуку та коефіцієнтом відношення 0.8. Така фільтрація дозволить не тільки більш точно розраховувати есенціальну матрицю, а й зменшити репроекційні похибки та більш точно відтворити реальний об'єкт/сцену за допомогою тривимірної реконструкції.

Останній результат на рис.3 не був включений у графік так як мав значення 1197.70px та данні з тріангуляції для цієї пари не були включені в загальний набір точок тривимірної моделі. Але загальна тенденція значно покращилась, данні

репроекційної помилки знаходяться в діапазоні від 0.015рх до 4.87рх і це суттєво менше в порівнянні зі звичайним методом SIFT.

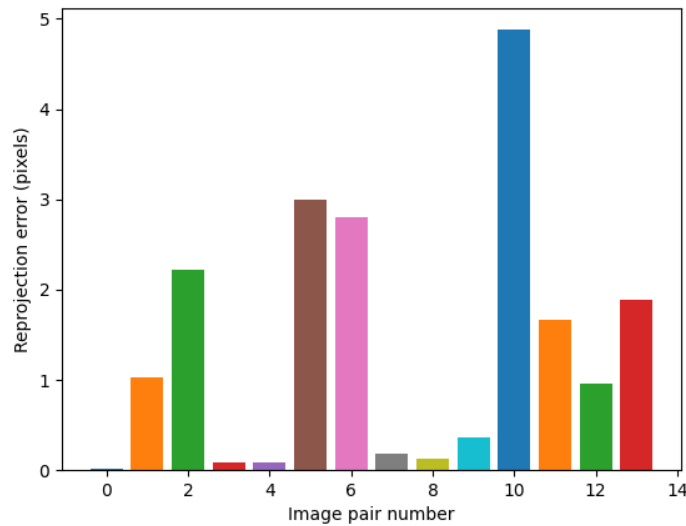


Рисунок 3 – Репроекційні похибки SFM алгоритму з використанням методу SIFT та поліпшеним обрахунком відповідності між ключовими точками.

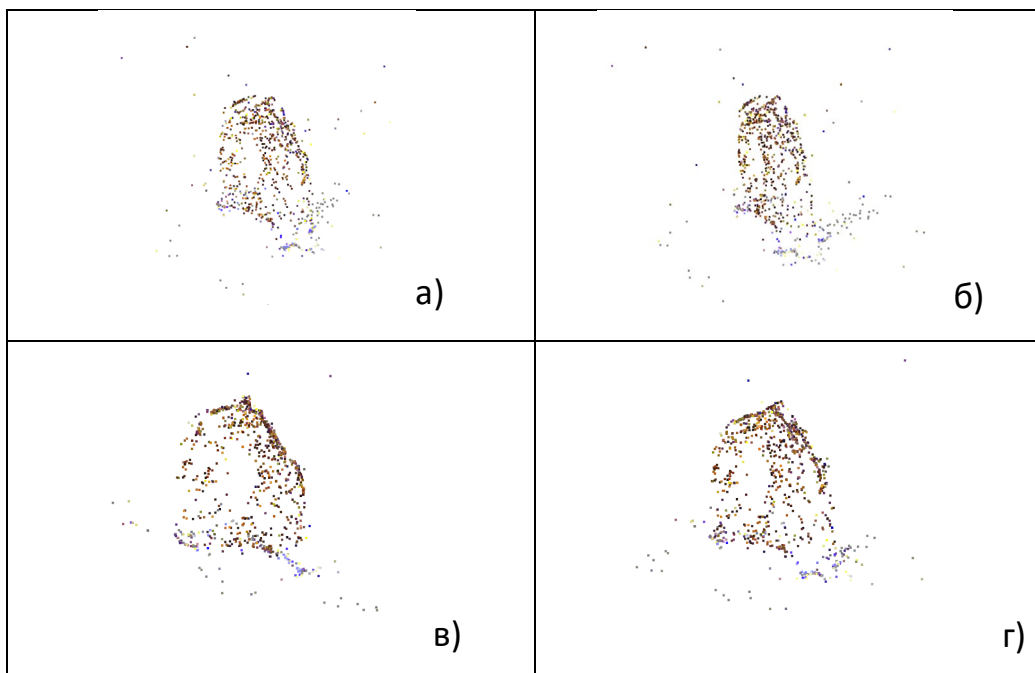


Рисунок 4 – Розріджена хмара точок на основі SFM алгоритму: а, б – використання стандартного методу знаходження відповідностей між парою зображень; в, г – використання методу з покращеною фільтрацією збігів між парою зображень.



За рис.4а та рис.4б видно, що точки які були отримані за допомогою стандартного методу визначення відповідностей між ключовими точками дають зашумлену розріджену хмару. Такі шуми пояснюються високими значеннями репроекційної похибки. На рис.4в та рис.4г демонструє більш чистішу розріджену хмару точок. Менше віддалених від основного об'єкта точок. Усі точки які відображають підлогу на якому стоїть зразок знаходяться на свої місцях. Лише пара точок віддалена від основного об'єкту. В цілому сцена демонструє коректність усього процесу триангуляції і визначає наскільки добре було відкалібровано камеру та наскільки точно визначено позиції камери $R|t$. Також дірки у розрідженій хмарі демонструють потребу в більшій кількості фотознімків саме цієї ділянки через слабе перекриття між зображеннями. Також на основі проведених досліджень запропонований метод додаткової перевірки правильності збігів продемонстрував кращі результати репроекційної похиби та на основі цього більш кращу триангуляцію. Тому такі дані алгоритми можна викристовувати як основу для SFM алгоритму, щоб в подальшому за допомогою MVS(multi view stereo) алгоритму отримати більш точніші данні щільної хмари точок.